

# CONVERGENCE OF REGULARIZED NONLOCAL INTERACTION ENERGIES

KATY CRAIG AND IHSAN TOPALOGLU

**ABSTRACT.** Inspired by numerical studies of the aggregation equation, we study the effect of regularization on nonlocal interaction energies. We consider energies defined via a repulsive-attractive interaction kernel, regularized by convolution with a mollifier. We prove that, with respect to the 2-Wasserstein metric, the regularized energies  $\Gamma$ -converge to the unregularized energy and minimizers converge to minimizers. We then apply our results to prove  $\Gamma$ -convergence of the gradient flows, when restricted to the space of measures with bounded density.

## 1. INTRODUCTION

We consider the nonlocal interaction energy

$$(1.1) \quad \mathbf{E}(\mu) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(x-y) d\mu(x) d\mu(y),$$

over the space  $\mathcal{P}_2(\mathbb{R}^d)$  of probability measures with finite second moment, where the pairwise interaction kernel  $K : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  is an even, locally integrable, and lower semicontinuous function. Depending on the choice of  $K$ , the asymptotic states of many physical and biological systems can be characterized as minimizers of this energy. Of particular interest are interaction kernels which are repulsive at short distances and attractive at long distances. Important examples of such repulsive-attractive kernels are Morse-type potentials

$$(1.2) \quad K(x) = C_r e^{-|x|/l_r} - C_a e^{-|x|/l_a}$$

with  $C_r/C_a < (l_r/l_a)^{-d}$ ,  $0 < l_r < l_a$ , and  $0 < C_a < C_r$ ,

and potentials in the power-law form

$$(1.3) \quad K(x) = |x|^q/q - |x|^p/p \quad \text{with} \quad -d < p < q,$$

which arise in models of granular media [9, 24, 25, 41], molecular self-assembly [33, 52, 62], biological swarming [10, 59], and the distributions of eigenvalues for Gaussian random matrices [27, 50].

Due to the competition between the repulsive and attractive terms, the minimization of these energies leads to complex equilibrium configurations [4, 11–14, 16, 32, 34–36, 39, 51, 57, 58]. In the case of repulsive-attractive power law kernels (1.3), local minimizers exhibit a variety

---

*Date:* October 12, 2015.

1991 *Mathematics Subject Classification.* 49J45, 82B21, 82B05, 35R09, 45K05.

*Key words and phrases.* global minimizers, pair potentials, aggregation, vortex blob method, Wasserstein metric, gradient flow,  $\Gamma$ -convergence, Coulomb repulsion.

KC's research is supported by U.S. National Science Foundation grant DMS 1401867 and the UC President's Postdoctoral Fellowship. IT's research is supported by a Fields–Ontario Postdoctoral Fellowship.

of qualitatively different patterns, from solid rings to broken, rounded triangles [40], and the dimension of the minimizers' support can be characterized in terms of the strength of the repulsive forces [6].

In addition to qualitative properties of local minimizers, there has also been significant interest in the existence and uniqueness of global minimizers. (Since (1.1) is translation invariant, minimizers are only unique up to translation.) Existence was recently established for a range of kernels, including (1.3), by Simione, Slepčev, and the second author [54] and by Cañizo, Carrillo, and Patacchini [18]. On the other hand, characterizing which kernels have a unique global minimizer is essentially open, due to lack of convexity. In the particular case of Coulomb repulsion and quadratic attraction (when the energy is convex with respect to  $H^{-1}$ ) the unique global minimizer is the indicator function on a ball (c.f. for example [21, 28]).

Alongside the static problem of characterizing the minimizers of the interaction energy, there has also been significant interest in the dynamic problem of understanding the behavior of systems as they evolve toward a local minimizer. Such systems arise in biological swarming [47, 48, 58], robotic swarming [29, 49], and assembly of viral capsid proteins [38], and they may be modeled as *gradient flows* of the energy with respect to the 2-*Wasserstein metric*. Formally, this corresponds to the nonlinear, nonlocal partial differential equation

$$(1.4) \quad \begin{cases} \rho_t + \nabla \cdot (v\rho) = 0 & \text{with } v = -2\nabla K * \rho, \\ \rho(x, 0) = \rho_0(x), \end{cases}$$

known as the *aggregation equation*. (We choose to put a factor of two in the velocity field instead of putting a factor of one half on the energy (1.1).) For semi-convex interaction kernels  $K$ , with up to a Lipschitz singularity, weak measure solutions to (1.4) exist for all time and are unique [1, 22]. If the kernels are sufficiently convex, solutions converge exponentially to a unique global minimizer [24]. However, for nonconvex kernels with merely integrable singularities, much less is known about the evolution of measure solutions and their asymptotic behavior. In the particular case of Coulomb repulsion and quadratic attraction, solutions with bounded, continuous initial data converge to the unique global minimizer algebraically in time [14].

Due to the analytical difficulties repulsive-attractive kernels present, both theoretical and applied work is often complemented by numerical simulations. The most common method for simulating (1.4) is a particle method, in which one approximates the initial data by a sum of Dirac masses,  $\rho_0 \approx \sum_{i=1}^N \delta_{x_i} m_i$ , where  $x_i$  is the location of the Dirac mass and  $m_i \geq 0$  is its weight. The corresponding solutions of (1.4) are formally of the form  $\rho \approx \sum_{i=1}^N \delta_{X_i(t)} m_i$ , where the trajectories of the Dirac masses  $X_i(t)$  satisfy the following finite system of ODEs:

$$(1.5) \quad \frac{d}{dt} X_i(t) = -2 \sum_{j=1}^N \nabla K(X_i(t) - X_j(t)) m_j, \quad X_i(0) = x_i.$$

For a range of interaction kernels, including repulsive-attractive power-laws (1.3) in the parameter regime  $2 - d < p \leq 2$ ,  $q > 0$ , Carrillo, Choi, and Hauray prove the convergence of the particle method solution  $\sum_{i=1}^N \delta_{X_i(t)} m_i$  to weak measure solutions of the aggregation equation [20].

In recent work, Bertozzi and the first author consider a modification of these types of particle methods [31], analogous to classical vortex blob methods from fluid dynamics (c.f. [3, 7, 8] and references therein). Specifically, since the gradient of the kernel  $\nabla K$  may be singular at the origin, they regularize the interaction kernel by convolution with a mollifier, leading to an

approximate solution of the form  $\sum_{i=1}^N \delta_{X_i^\epsilon(t)} m_i$ , where  $X_i^\epsilon(t)$  satisfies (1.5) with  $\nabla K$  replaced by  $\nabla(K * \varphi_\epsilon)$ . By regularizing the interaction kernel  $K$ , the authors extend particle methods to a wide range of interaction kernels, including the power-law kernels (1.3) for  $2 - d \leq p \leq q$ , and obtain quantitative rates of convergence to classical solutions. While this convergence result is limited to bounded time intervals, numerical results indicate that regularizing the energy in this way may also be useful in studying asymptotic behavior (see Figure 1).

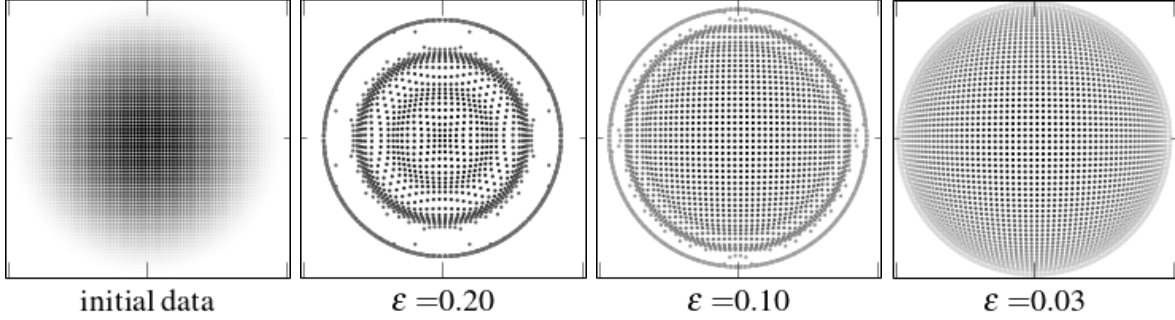


FIGURE 1. Numerical solutions of (1.4) at  $t = 100$  for Coulomb repulsion and quadratic attraction in two dimensions. The initial data is the function  $\rho_0(x, y) = C(1 - x^2 - y^2)_+^2$ , with  $C$  chosen so  $\rho_0$  has mass one. Solutions are generated by the blob method [31] with  $\epsilon = 0.20, 0.10$ , and  $0.03$  on the square  $[-1, 1] \times [-1, 1]$ , which is discretized on a  $80 \times 80$  grid. As  $\epsilon \rightarrow 0$ , the support of the numerical solution spreads evenly throughout the ball. This suggests that, in the limit, the solution converges to the indicator function on the ball, the unique steady state of the PDE.

Inspired by these results, we consider a regularization of the nonlocal interaction energy analogous to Bertozzi and the first author's blob method. Specifically, given a smooth, radial, rapidly decreasing mollifier  $\varphi$ , we define the regularized kernel

$$K_\epsilon := \varphi_\epsilon * K * \varphi_\epsilon,$$

with  $\varphi_\epsilon(x) := \epsilon^{-d} \varphi(x/\epsilon)$ , and consider the regularized interaction energy

$$(1.6) \quad E_\epsilon(\mu) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_\epsilon(x - y) d\mu(x) d\mu(y).$$

We show that, for a range of repulsive-attractive kernels, including repulsive-attractive power-law kernels (1.3) with  $2 - d \leq p < 0 < q \leq 2$ , the regularized energies  $E_\epsilon$   $\Gamma$ -converge to  $E$  with respect to the 2-Wasserstein metric. Then, via a compactness argument, we show that, up to a subsequence, the minimizers of  $E_\epsilon$  converge to a minimizer of  $E$  as  $\epsilon \rightarrow 0$ . This provides a method of approximating minimizers of  $E$  by minimizers of energies with superior convexity and regularity properties (see Remarks 2.7 and 2.8). It also demonstrates a type of continuity among minimizers of  $E_\epsilon$ , in spite of their vastly different regularity properties, as illustrated in Figure 1. Finally, it imparts further theoretical justification for the success of the Bertozzi and the first author's numerical blob method. Our work builds on previous work of Fellner and Raoul [34], who considered regularization of the Coulomb repulsion in one dimension, showing

that, in the presence of a confining potential, Dirac-type stationary states converge to a unique  $L^\infty$  stationary state.

We conclude by considering the convergence of the corresponding gradient flows for repulsive-attractive power-law kernels (1.3) in the regime  $2 - d \leq p < 0 < q \leq 2$ . Indeed, using the abstract scheme of Serfaty [53], we show that as  $\epsilon \rightarrow 0$  gradient flows of  $E_\epsilon$  that are bounded in  $L^\infty(\mathbb{R}^d)$   $\Gamma$ -converge to a generalized notion of gradient flow for  $E$ , that is to its *curve of maximal slope* on the metric space of probability measures with bounded density. This provides a link between the well-understood case of gradient flows of semi-convex energies  $E_\epsilon$  and the curve of maximal slope of the unregularized energy  $E$ . It also provides a first step in understanding the connection between the gradient flows of unregularized and nonconvex interaction energies  $E$  and the aggregation equation, via a singular perturbation approach (see Remark 5.2).

These results provide many directions for future work. A natural extension would be to consider the effect of simultaneously discretizing and regularizing the energy and to study the distinguished limits as the discretization and regularization are removed. Another interesting direction would be to seek quantitative estimates on the rate of convergence of minimizers and gradient flows as a regularization is removed. Our present work strongly leverages compactness arguments, and we believe that a very different approach would be needed in order to provide quantitative estimates.

*Remark 1.1* (Power-law kernels and generalizations). In sections 2 and 3, we choose to state our results for interaction kernels in the power-law form (1.3), not only because these kernels have attracted much interest in connection with the aggregation equation (cf. [15, 32, 35, 36, 63]) and nonlocal interaction energies (cf. [19, 21, 28]), but also because these kernels capture the crucial behavior of short-range and long-range interactions, i.e., at the origin and at infinity. Perhaps one of the most important properties of such kernels, which we exploit in our analysis, is that the attractive and repulsive parts of the kernel are given separately by relatively simple potentials. This allows us to quantitatively compare the attractive and repulsive parts of the energy. Indeed, it is due to the comparison (or rather competition) between these two terms that minimization of such energies exhibit energy-driven pattern formation.

Our analysis depends mostly on the behavior of  $K$  for very large and very small values of  $|x|$ ; hence, there are many possible ways of extending our results in sections 2 and 3 to a wider class of kernels. In section 4, we provide general conditions on the regularity, growth at infinity, and behavior around singularities of  $K$  for which our results continue to hold. In particular, our results hold in any dimension  $d \geq 1$  for kernels that exhibit power-law repulsion at the origin and power-law attraction at infinity, including logarithmic repulsion in  $d \geq 2$ .

*Remark 1.2* (Choices of parameters in power-law kernels). In sections 2 and 3, we restrict our attention to repulsive-attractive power law kernels satisfying (1.3) with  $2 - d \leq p < 0 < q \leq 2$  in dimension  $d \geq 3$ . We require  $q \leq 2$  so that the attractive term and its regularization are functions with at most quadratic growth and their convolution with  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  is finitely valued (see Remark 2.9). We require  $q > 0$  to obtain compactness of the sequence of minimizers corresponding to the regularized energies  $E_\epsilon$  (see Remark 3.8). We take  $p < 0$  since our results are trivially true for nonsingular attractive-repulsive power-law kernels (see section 4). On the other hand, we take  $2 - d \leq p$  in order to ensure monotonicity of the repulsive part of the regularized kernels with respect to  $\epsilon$ , when the mollifier is the heat kernel (see Proposition 2.6).

Our paper is organized as follows. In section 2, we recall fundamental results on the 2-Wasserstein metric and the regularization of probability measures. We use these to establish the above mentioned monotonicity property of the regularized energies (1.6) in the regularization parameter  $\epsilon > 0$ . In section 3, we prove the  $\Gamma$ -convergence of  $E_\epsilon$  to  $E$  and the corresponding convergence of minimizers. In section 4, we extend these results to more general interaction kernels. Finally, in section 5, we introduce background on curves of maximal slope and prove the  $\Gamma$ -convergence of these generalized gradient flows for repulsive-attractive power-law kernels restricted to bounded densities.

Throughout, we denote by  $L^p(\mathbb{R}^d)$  the space of functions whose  $p$ -th power is integrable with respect to the Lebesgue measure on  $\mathbb{R}^d$  and by  $L^p(\mu)$  those functions whose  $p$ -th power is integrable with respect to the probability measure  $\mu \in \mathcal{P}(\mathbb{R}^d)$ .  $C(\mathbb{R}^d)$  denotes the space of continuous functions on  $\mathbb{R}^d$ ,  $C^\infty(\mathbb{R}^d)$  denotes the space of smooth functions on  $\mathbb{R}^d$ , and  $B_r(x)$  denotes the open ball of radius  $r$  centered at  $x \in \mathbb{R}^d$ . All constants which appear in the text are positive and may change from line to line. We express the dependence of constants to parameters by a subscript such as  $C_{d,p}$  when necessary.

## 2. PRELIMINARIES

**2.1. The 2-Wasserstein metric and regularization of probability measures.** We consider the energies  $E_\epsilon$  and  $E$  over the space

$$\mathcal{P}_2(\mathbb{R}^d) := \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|^2 d\mu(x) < +\infty \right\}$$

of probability measures with finite second moment. We endow this space with the 2-Wasserstein metric, which we recall briefly now, along with some of its basic properties. For further background, we refer the reader to the books by Ambrosio, Gigli and Savaré [1] and Villani [61].

The 2-Wasserstein distance between  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  is

$$(2.1) \quad d_W(\mu, \nu) := \left( \min \left\{ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |x - y|^2 d\gamma(x, y) : \gamma \in \mathcal{C}(\mu, \nu) \right\} \right)^{1/2},$$

where  $\mathcal{C}(\mu, \nu)$  is the set of transport plans between  $\mu$  and  $\nu$ ,

$$\mathcal{C}(\mu, \nu) := \left\{ \gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) : (\pi_1)_\# \gamma = \mu \quad \text{and} \quad (\pi_2)_\# \gamma = \nu \right\}.$$

Here  $\pi_1, \pi_2$  denote the projections  $\pi_1(x, y) = x$  and  $\pi_2(x, y) = y$ . For  $i = 1, 2$ ,  $(\pi_i)_\# \gamma$  denotes the pushforward of  $\gamma$  defined by  $(\pi_i)_\# \gamma(U) := \gamma(\pi_i^{-1}(U))$  for any measurable set  $U \subset \mathbb{R}^d$ .

The minimization problem (2.1) admits a solution, i.e., there exists an optimal transport plan  $\gamma_0 \in \mathcal{C}_0(\mu, \nu)$  so that

$$d_W^2(\mu, \nu) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |x - y|^2 d\gamma_0(x, y).$$

Moreover,  $(\mathcal{P}_2(\mathbb{R}^d), d_W)$  is a complete and separable metric space, and convergence in  $(\mathcal{P}_2(\mathbb{R}^d), d_W)$  can be characterized as follows:

$$\begin{aligned} d_W(\mu_n, \mu) \rightarrow 0 &\iff \mu_n \rightarrow \mu \text{ weak-}^* \text{ in } \mathcal{P}(\mathbb{R}^d) \text{ and } \int_{\mathbb{R}^d} |x|^2 d\mu_n(x) \rightarrow \int_{\mathbb{R}^d} |x|^2 d\mu(x), \\ &\iff \int_{\mathbb{R}^d} f(x) d\mu_n(x) \rightarrow \int_{\mathbb{R}^d} f(x) d\mu(x), \\ &\quad \forall f \in C(\mathbb{R}^d) \text{ such that } |f(x)| \leq C(1 + |x - x_0|^2). \end{aligned}$$

We will refer to functions satisfying  $|f(x)| \leq C(1 + |x - x_0|^2)$ , for some  $C > 0$  and  $x_0 \in \mathbb{R}^d$ , as functions with *at most quadratic growth*.

In addition to regularizing our energy functionals by convolution with a mollifier, we will also regularize our measures. We recall the definition of the convolution of a measure.

**Definition 2.1.** For  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  and  $\varphi$  continuous with at most quadratic growth,  $\mu * \varphi$  is defined by

$$(2.2) \quad \int_{\mathbb{R}^d} f(x) d(\mu * \varphi)(x) = \int_{\mathbb{R}^d} f * \varphi(y) d\mu(y),$$

for all bounded measurable functions  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ .

*Remark 2.2.* Note that  $\mu * \varphi$  is absolutely continuous with respect to Lebesgue measure and  $d(\mu * \varphi) = \mu * \varphi(x) dx$ , where  $\mu * \varphi(x) = \int_{\mathbb{R}^d} \varphi(x - y) d\mu(y)$ . If, in addition,  $\varphi(x) \in L^\infty(\mathbb{R}^d)$ , then  $\mu * \varphi(x) \in L^\infty(\mathbb{R}^d)$  and (2.2) holds for all  $f \in L^1(\mathbb{R}^d)$ .

We recall the following lemma on the approximation of measures by convolution.

**Lemma 2.3** (cf. [1, Lemma 7.1.10]). Fix  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  and a mollifier  $\varphi \in C^\infty(\mathbb{R}^d)$  with finite second moment,  $\varphi \geq 0$ , and  $\int_{\mathbb{R}^d} \varphi(x) dx = 1$ . Then for  $\varphi_\epsilon(x) := \epsilon^{-d} \varphi(x/\epsilon)$ ,  $\mu * \varphi_\epsilon \in \mathcal{P}_2(\mathbb{R}^d)$  and

$$(2.3) \quad d_W(\mu * \varphi_\epsilon, \mu) \leq \epsilon \left( \int_{\mathbb{R}^d} |x|^2 \varphi(x) dx \right)^{1/2}.$$

**2.2. Regularization of energies.** In what follows, we consider mollifiers which satisfy the following assumptions:

- (M1)  $\varphi \in C^\infty(\mathbb{R}^d)$  and  $\varphi(x) = \varphi(|x|) \geq 0$  for all  $x \in \mathbb{R}^d$ ,
- (M2)  $\int_{\mathbb{R}^d} \varphi(x) dx = 1$ ,
- (M3)  $\varphi(x) \leq C|x|^{-l}$  for some  $l \geq \max\{2d + 1, 4\}$  and  $C > 0$ .

*Remark 2.4* (Finite second moment of mollifiers). Any mollifier  $\varphi$  satisfying the assumptions (M1)–(M3) has finite second moment, i.e.,  $\int_{\mathbb{R}^d} |x|^2 \varphi(x) dx < +\infty$ .

As a consequence of the previous results on regularization of measures in the Wasserstein metric, we have the following lemma which relates  $E_\epsilon$  to  $E$ .

**Lemma 2.5.** Consider an interaction energy  $E(\mu) = \int \int_{\mathbb{R}^d \times \mathbb{R}^d} K(x - y) d\mu(x) d\mu(y)$ , where  $K(x)$  is an even, nonnegative, locally integrable function that has at most quadratic growth and is continuous outside of some ball. Given a mollifier  $\varphi$  satisfying (M1)–(M3), then for all  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$

$$E_\epsilon(\mu) = E(\mu * \varphi_\epsilon).$$



*Proof.* Define the right shift operator  $\tau_y : f(x) \mapsto f(x - y)$ . For any  $f$  and  $g$  even,

$$\begin{aligned}\tau_y[f * g](x) &= f * g(x - y) = \int_{\mathbb{R}^d} f(x - y - z)g(z) dz \\ &= \int_{\mathbb{R}^d} \tau_y[f](x - z)g(z) dz = \tau_y[f] * g(x).\end{aligned}$$

Furthermore, since  $f * g(x - y) = f * g(y - x)$ , we also have  $\tau_y[f * g](x) = \tau_x[f] * g(y)$ . Since  $\varphi \in L^\infty(\mathbb{R}^d)$ , (2.2) holds for all continuous  $f$  of at most quadratic growth and for all  $f \in L^1(\mathbb{R}^d)$ . As we may decompose  $K$  into the sum of a continuous function of at most quadratic growth (away from the origin) and an integrable component (near the origin), the result then follows:

$$\begin{aligned}\mathbf{E}_\epsilon(\mu) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi_\epsilon * K * \varphi_\epsilon(x - y) d\mu(x) d\mu(y) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \tau_y[\varphi_\epsilon * K] * \varphi_\epsilon(x) d\mu(x) d\mu(y), \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi_\epsilon * \tau_y[K] * \varphi_\epsilon(x) d\mu(x) d\mu(y) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi_\epsilon * \tau_y[K](x) d[\mu * \varphi_\epsilon](x) d\mu(y), \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi_\epsilon * \tau_x[K](y) d[\mu * \varphi_\epsilon](x) d\mu(y) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(x - y) d[\mu * \varphi_\epsilon](y) d[\mu * \varphi_\epsilon](x), \\ &= \mathbf{E}(\mu * \varphi_\epsilon).\end{aligned}$$

□

For general interaction kernels and mollifiers, there is no uniform relation between the size of  $\mathbf{E}_\epsilon(\mu)$  compared to  $\mathbf{E}(\mu)$ . Consequently, in this section and section 3, we consider interaction energies of the following form:

$$(E1) \quad \mathbf{E}(\mu) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} K(x - y) d\mu(x) d\mu(y) \text{ for } K(x) = |x|^q/q - |x|^p/p,$$

$$(E2) \quad 2 - d \leq p < 0 < q \leq 2.$$

If we mollify energies satisfying (E1)–(E2) via the heat kernel, we obtain monotonicity in  $\epsilon$  of the repulsive part of the interaction energy. This generalizes a property used by Blanchet, Carlen, and Carrillo [17] for the Newtonian potential. In section 4, we describe how to extend these results to more general interaction energies.

Let  $K^a := \frac{1}{q}|x|^q$  denote the attractive part of the interaction kernel and  $K^r := -\frac{1}{p}|x|^p$  the repulsive part. Similarly, let  $\mathbf{E}^a$  denote the attractive part of the energy and  $\mathbf{E}^r$  denote the repulsive part,

$$(2.4) \quad \mathbf{E}^a(\mu) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^a(x - y) d\mu(x) d\mu(y), \quad \mathbf{E}^r(\mu) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^r(x - y) d\mu(x) d\mu(y).$$

**Proposition 2.6** (Monotonicity of  $\mathbf{E}_\epsilon^r$ ). *Suppose the energy  $\mathbf{E}$  satisfies (E1)–(E2) and the regularization mollifier  $\psi_\epsilon$  is given by the heat kernel. Then for all  $\epsilon_1 \geq \epsilon_2 > 0$  and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$*

$$\mathbf{E}_{\epsilon_1}^r(\mu) \leq \mathbf{E}_{\epsilon_2}^r(\mu).$$

*Proof.* Since  $\psi_\epsilon(x) = \psi(\epsilon, x) = (4\pi\epsilon)^{-d/2} e^{-|x|^2/4\epsilon}$ ,  $\partial_\epsilon \psi = \Delta_x \psi$  for all  $\epsilon > 0$  and  $x \in \mathbb{R}^d$ . In the parameter regime  $2 - d < p < 0$ , the Riesz potential

$$I_{d+p}(f)(x) = C_{d,p} \int_{\mathbb{R}^d} f(y) |x - y|^p dy = -C_{d,p} p K^r * f(x)$$

satisfies the identity

$$I_{d+p}(\Delta f) = \Delta I_{d+p}(f) = -I_{d+p-2}(f)$$

for any Schwartz class function  $f$  (see e.g. [55]). Moreover,  $I_{d+p}(f) \geq 0$  for any  $f \geq 0$ .

Defining  $K_\epsilon^r(x) = \psi_\epsilon * K^r * \psi_\epsilon$  and using the above identity we obtain

$$\begin{aligned} \frac{\partial}{\partial \epsilon} K_\epsilon^r(x) &= 2\psi_\epsilon * K^r * (\partial_\epsilon \psi_\epsilon) = 2\psi_\epsilon * K^r * (\Delta \psi_\epsilon) = -(p C_{d,p})^{-1} \psi_\epsilon * I_{d+p}(\Delta \psi_\epsilon) \\ &= (p C_{d,p})^{-1} \psi_\epsilon * I_{p+d-2}(\psi_\epsilon) \leq 0 \end{aligned}$$

since  $p < 0$  and  $\psi_\epsilon \geq 0$ . Thus,  $K_\epsilon^r$  is monotonically decreasing in  $\epsilon$ .

When  $p = 2 - d$ ,  $I_{d+p}(f)$  is the Newtonian potential and satisfies the identity  $I_{d+p}(\Delta f) = \Delta I_{d+p}(f) = -f$ . Again, differentiating  $K_\epsilon^r$  with respect to  $\epsilon$  yields the desired result. Therefore, by definition of  $\mathbf{E}^r$  (2.4), we conclude monotonicity in  $\epsilon$ .  $\square$

We close this section with brief remarks on the convexity and differentiability of regularized energies.

*Remark 2.7* ( $\lambda_\epsilon$ -convexity of  $\mathbf{E}_\epsilon$ ). Recall that a function  $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  is  $\lambda$ -convex if  $f(x) - \frac{\lambda}{2}|x|^2$  is convex for some  $\lambda \leq 0$ . If  $f$  is twice continuously differentiable, this is equivalent to  $D^2 f(x) \geq \lambda I$  for all  $x \in \mathbb{R}^d$ . Consequently, if  $\varphi$  satisfies (M1)–(M3) and  $\mathbf{E}$  satisfies (E1)–(E2), then both the regularized potentials  $K_\epsilon$  and their opposites  $-K_\epsilon$  are  $\lambda_\epsilon$ -convex with  $\lambda_\epsilon = -C_\varphi \epsilon^{-d}$ , where  $C_\varphi$  is a positive constant depending on  $\varphi$  and  $\epsilon$  is sufficiently small. For an interaction energy of the form (1.1),  $\lambda$ -convexity of the kernel  $K$  ensures  $\lambda$ -convexity of the energy  $\mathbf{E}$  with respect to the Wasserstein metric [22]. Thus,  $\mathbf{E}_\epsilon$  and  $-\mathbf{E}_\epsilon$  are both  $\lambda_\epsilon$ -convex.

*Remark 2.8* (Differentiability of  $\mathbf{E}_\epsilon$ ). Given a functional  $\mathbf{F} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$  and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  so that  $\mathbf{F}(\mu) < +\infty$ , the *metric local slope* of  $\mathbf{F}$  at  $\mu$  is

$$|\partial \mathbf{F}|(\mu) := \limsup_{\nu \rightarrow \mu} \frac{(\mathbf{F}(\mu) - \mathbf{F}(\nu))_+}{d_W(\mu, \nu)}.$$

If  $\varphi$  satisfies (M1)–(M3) and  $\mathbf{E}$  satisfies (E1)–(E2), the metric local slope for  $\mathbf{E}_\epsilon$  is well-defined (cf. [22, Proposition 2.2], [1, Lemma 10.1.5]) and

$$|\partial \mathbf{E}_\epsilon|(\mu) = 2 \|\nabla K_\epsilon * \mu\|_{L^2(\mu)}.$$

Furthermore, for  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ , we have  $2 \|\nabla K_\epsilon * \mu\|_{L^2(\mu)} \leq C_{\varphi, \mu} \epsilon^{1-d}$  for  $\epsilon$  sufficiently small. To see this, let  $0 \leq \eta(x) \leq 1$  be a smooth function satisfying  $\eta(x) = 0$  if  $|x| \leq 1$  and  $\eta(x) = 1$  if  $|x| \geq 2$ . Using  $\eta$ , we decompose  $\nabla K_\epsilon$  into its singular and nonsingular components,  $\nabla K_\epsilon = \nabla K_\epsilon(1 - \eta) + \nabla K_\epsilon \eta = \nabla K_\epsilon^s + \nabla K_\epsilon^n$ . By linearity of convolution and the triangle inequality for  $L^2(\mu)$ , it suffices to estimate  $\|\nabla K_\epsilon^s * \mu\|_{L^2(\mu)}$  and  $\|\nabla K_\epsilon^n * \mu\|_{L^2(\mu)}$ . The former is bounded by  $\|\nabla K_\epsilon^s * \mu\|_{L^\infty(d\mu)} \leq \sup_{x \in \mathbb{R}^d} |\nabla K_\epsilon^s(x)| \leq C_\varphi \epsilon^{1-d}$ . To bound the latter, we apply Minkowski's integral inequality, the fact that  $\nabla K_\epsilon^n$  has at most linear growth, and the fact that  $\mu$  has finite second moment to obtain

$$\|\nabla K_\epsilon^n * \mu\|_{L^2(\mu)} \leq \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} |\nabla K_\epsilon^n(x - y)|^2 d\mu(y) \right)^{1/2} d\mu(x) \leq C_{\varphi, \mu}.$$



*Remark 2.9* (Attraction powers  $q > 2$ ). One possible way to extend our results to cover attraction powers  $q > 2$  would be to consider the energy  $E$  over the space  $\mathcal{P}_q(\mathbb{R}^d)$  of probability measures with finite moments up to order  $q$ , endowed with the  $q$ -Wasserstein distance

$$d_q(\mu, \nu) := \left( \min \left\{ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |x - y|^q d\gamma(x, y) : \gamma \in \mathcal{C}(\mu, \nu) \right\} \right)^{1/q},$$

where  $\mathcal{C}(\mu, \nu)$  is the set of transport plans between  $\mu$  and  $\nu$ . Alternatively, to cover all  $q > 0$ , one could consider the  $\infty$ -Wasserstein metric

$$d_\infty(\mu, \nu) := \inf_{\gamma \in \mathcal{C}(\mu, \nu)} \sup_{(x, y) \in \text{supp}(\gamma)} |x - y|$$

over the space  $\mathcal{P}_\infty(\mathbb{R}^d)$  of probability measures with finite moments of all orders. The space  $(\mathcal{P}_\infty(\mathbb{R}^d), d_\infty)$  is a complete metric space [37] and has been used in the study of local minimizers for several variational problems, including nonlocal repulsive-attractive energies [6, 21, 46].

The main difficulty in extending our results to  $\mathcal{P}_q(\mathbb{R}^d)$  or  $\mathcal{P}_\infty(\mathbb{R}^d)$  lies in understanding the appropriate generalization of the  $\lambda$ -convexity to these distances and, in particular, the appropriate analogue of the HWI inequality, which plays a key role in the proof of Theorem 3.4. We leave this to future work.

### 3. CONVERGENCE OF REGULARIZED ENERGIES AND MINIMIZERS

We now turn to the proof that, up to a subsequence, minimizers of the regularized energies  $E_\epsilon$  converge to a minimizer  $\mu$  of  $E$  with respect to  $d_W$ . We establish this by first proving a  $\Gamma$ -convergence result for the sequence of energies  $E_\epsilon$  and then obtaining compactness for any sequence  $\{\mu_\epsilon\}_{\epsilon>0}$  when  $E_\epsilon(\mu_\epsilon)$  is uniformly bounded. For the latter, we use Lions' result on concentration compactness, which we recall for the readers' convenience.

**Lemma 3.1** (Concentration compactness lemma for measures (cf. [43], [56, Section 4.3])). *Let  $\{\mu_n\}_{n \in \mathbb{N}}$  be a sequence of probability measures on  $\mathbb{R}^d$ . Then there exists a subsequence  $\{\mu_{n_k}\}_{k \in \mathbb{N}}$  satisfying one of the three following possibilities:*

- (i) (tightness up to translation) *There exists a sequence  $\{y_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^d$  such that for all  $\epsilon > 0$  there exists  $R > 0$  with the property that*

$$\int_{B_R(y_k)} d\mu_{n_k}(x) \geq 1 - \epsilon \quad \text{for all } k.$$

- (ii) (vanishing)  *$\lim_{k \rightarrow \infty} \sup_{y \in \mathbb{R}^d} \int_{B_R(y)} d\mu_{n_k}(x) = 0$ , for all  $R > 0$ ;*

- (iii) (dichotomy) *There exists  $\alpha \in (0, 1)$  such that for all  $\epsilon > 0$ , there exist a number  $R > 0$  and a sequence  $\{x_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^d$  with the following property:*

*Given any  $R' > R$  there are nonnegative measures  $\mu_k^1$  and  $\mu_k^2$  such that*

$$0 \leq \mu_k^1 + \mu_k^2 \leq \mu_{n_k}, \quad \text{supp}(\mu_k^1) \subset B_R(x_k), \quad \text{supp}(\mu_k^2) \subset \mathbb{R}^d \setminus B_{R'}(x_k),$$

$$\limsup_{k \rightarrow \infty} \left( \left| \alpha - \int_{\mathbb{R}^d} d\mu_k^1(x) \right| + \left| (1 - \alpha) - \int_{\mathbb{R}^d} d\mu_k^2(x) \right| \right) \leq \epsilon.$$

We begin our proof of the convergence of minimizers with the observation that minimizers exist for both the regularized and unregularized energies.

**Proposition 3.2** (Existence of minimizers in  $\mathcal{P}_2(\mathbb{R}^d)$ ). *Suppose the energy  $E$  satisfies (E1)–(E2) and the mollifier  $\varphi$  satisfies (M1)–(M3). Then both  $E$  and  $E_\epsilon$  admit a minimizer in  $\mathcal{P}_2(\mathbb{R}^d)$ , for  $\epsilon > 0$  sufficiently small.*

*Proof.* First note that, in the parameter regime  $2-d < p < 0 < q \leq 2$ , the interaction potential  $K$  and its regularization  $K_\epsilon$  are locally integrable and lower semicontinuous.

We now show that  $K$  and  $K_\epsilon$  are strictly increasing in each coordinate for  $x_i$  sufficiently large. For  $K$  given as in (E1)–(E2) and  $x \neq 0$ ,

$$\partial_{x_i} K(x) = x_i |x|^{q-2} - x_i |x|^{p-2} = x_i |x|^{q-2} (1 - |x|^{p-q}).$$

Consequently, if  $x_i > 1$ ,  $\partial_{x_i} K(x)$  is positive.

Now we consider  $K_\epsilon$ . It suffices to show that if  $x_i > 2$ , then

$$(3.1) \quad |\partial_{x_i} K_\epsilon(x) - \partial_{x_i} K(x)| \leq C\epsilon |x|^{q-2}$$

for some constant  $C > 0$ . Then for  $x_i > 2$ , we have  $(1 - |x|^{p-q}) > \tilde{C}_{p,q} > 0$  and

$$\begin{aligned} \partial_{x_i} K_\epsilon(x) &= \partial_{x_i} K(x) + \partial_{x_i} K_\epsilon(x) - \partial_{x_i} K(x) \geq x_i |x|^{q-2} (1 - |x|^{p-q}) - C\epsilon |x|^{q-2} \\ &\geq 2|x|^{q-2} \tilde{C}_{p,q} - C\epsilon |x|^{q-2} = |x|^{q-2} (2\tilde{C}_{p,q} - C\epsilon). \end{aligned}$$

Choosing  $\epsilon$  sufficiently small, this is nonnegative.

We now prove inequality (3.1). First, we use a cutoff function to rewrite  $\partial_{x_i} K(x)$  as the sum of a compactly supported singular function  $\partial_{x_i} K^s$  and a continuously differentiable function  $\partial_{x_i} K^n$ . Let  $0 \leq \eta \leq 1$  be a smooth function satisfying  $\eta(x) \equiv 0$  for  $|x| < 1/4$  and  $\eta(x) \equiv 1$  for  $|x| > 1/2$ . Write  $\partial_{x_i} K = (1 - \eta)\partial_{x_i} K + \eta\partial_{x_i} K =: \partial_{x_i} K^s + \partial_{x_i} K^n$ . Then, for  $\Phi_\epsilon(y) := \int_{\mathbb{R}^d} \varphi_\epsilon(y - z)\varphi_\epsilon(z) dz$ ,

$$\begin{aligned} |\partial_{x_i} K_\epsilon(x) - \partial_{x_i} K(x)| &= \left| \int_{\mathbb{R}^d} (\partial_{x_i} K(x - y) - \partial_{x_i} K(x)) \Phi_\epsilon(y) dy \right| \\ &\leq \left| \int_{\mathbb{R}^d} (\partial_{x_i} K^s(x - y) - \partial_{x_i} K^s(x)) \Phi_\epsilon(y) dy \right| \\ &\quad + \left| \int_{\mathbb{R}^d} (\partial_{x_i} K^n(x - y) - \partial_{x_i} K^n(x)) \Phi_\epsilon(y) dy \right| \\ &=: I_1 + I_2. \end{aligned}$$

The autocorrelation function  $\Phi_\epsilon$  satisfies  $\Phi_\epsilon(y) = \epsilon^{-d}\Phi(y/\epsilon)$  for all  $y \in \mathbb{R}^d$  and has the same decay property (M3) as  $\varphi_\epsilon$ . To see this, suppose  $|y| = 3R$  for some  $R > 0$ . Then for any  $z \in \mathbb{R}^d$  we have  $|z| > R$  or  $|y - z| > R$ ; hence,

$$\begin{aligned} |\Phi_\epsilon(y)| &= \frac{1}{\epsilon^{2d}} \int_{\mathbb{R}^d} \varphi\left(\frac{y-z}{\epsilon}\right) \varphi\left(\frac{z}{\epsilon}\right) dz \\ &\leq \frac{1}{\epsilon^{2d}} \left( \int_{|z|>R} \varphi\left(\frac{y-z}{\epsilon}\right) \varphi\left(\frac{z}{\epsilon}\right) dz + \int_{|y-z|>R} \varphi\left(\frac{y-z}{\epsilon}\right) \varphi\left(\frac{z}{\epsilon}\right) dz \right) \\ &\leq C\epsilon^{l-2d} R^{-l} = C\epsilon^{l-2d} |y|^{-l}. \end{aligned}$$

First, we estimate  $I_1$ . If  $x_i > 2$ , then  $|x| > 2$ , so  $\partial_{x_i} K^s(x) = 0$ . Consequently,

$$I_1 = \left| \int_{|x-y|<1} \partial_{x_i} K^s(x-y) \Phi_\epsilon(y) dy \right|.$$

Since  $|x-y| < 1$  and  $|x| > 2$ ,

$$|y| \geq |x| - |y-x| \geq |x| - 1 \geq |x|/2.$$

Consequently,

$$|\Phi_\epsilon(y)| \leq C \epsilon^{l-2d} |y|^{-l} \leq C \epsilon |y|^{-l} \leq C 2^l \epsilon |x|^{-l}$$

for  $0 < \epsilon < 1$ . Since  $\partial_{x_i} K^s$  is locally integrable and  $2 - q < 2d + 1 \leq l$ , there exists  $C > 0$  so that for  $|x| > 2$ ,

$$I_1 \leq C \epsilon |x|^{-l} \leq C \epsilon |x|^{q-2}.$$

Now we estimate  $I_2$ . Since  $\partial_{x_i} K^n$  is continuously differentiable,

$$\begin{aligned} I_2 &= \left| \int_{\mathbb{R}^d} \int_0^1 \frac{d}{d\alpha} \partial_{x_i} K^n(x - \alpha y) d\alpha \Phi_\epsilon(y) dy \right| \\ &= \left| \int_{\mathbb{R}^d} \int_0^1 \langle \nabla \partial_{x_i} K^n(x - \alpha y), -y \rangle d\alpha \Phi_\epsilon(y) dy \right| \\ &\leq \int_{|y| \leq |x|/2} \int_0^1 |\nabla \partial_{x_i} K^n(x - \alpha y)| |y| d\alpha \Phi_\epsilon(y) dy \\ &\quad + \int_{|y| > |x|/2} \int_0^1 |\nabla \partial_{x_i} K^n(x - \alpha y)| |y| d\alpha \Phi_\epsilon(y) dy. \end{aligned}$$

For  $|y| \leq |x|/2$ ,  $|x - \alpha y| \geq |x|/2 > 1$  for all  $\alpha \in [0, 1]$ . Thus,  $\nabla \eta(x - \alpha y) \equiv 0$ , and

$$|\nabla \partial_{x_i} K^n(x - \alpha y)| \leq |(\nabla \eta(x - \alpha y)) \partial_{x_i} K(x - \alpha y)| + |\eta(x - \alpha y) \nabla \partial_{x_i} K(x - \alpha y)| \leq C |x|^{q-2}.$$

On the other hand, for all  $y \in \mathbb{R}^d$ ,  $|\nabla \partial_{x_i} K^n(x - \alpha y)| \leq C$ . Therefore,

$$\begin{aligned} I_2 &\leq C |x|^{q-2} \int_{|y| < |x|/2} |y| \Phi_\epsilon(y) dy + C \int_{|y| > |x|/2} |y| \Phi_\epsilon(y) dy \\ &= C \epsilon |x|^{q-2} \int_{|z| < |x|/(2\epsilon)} |z| \Phi(z) dz + C \epsilon \int_{|z| > |x|/(2\epsilon)} |z| \Phi(z) dz. \end{aligned}$$

The first integral is bounded by a constant since  $\Phi$  satisfies decay property (M3), hence, has finite first moment. For the second integral, we use that  $|z\Phi(z)| \leq C|z|^{-l+1}$  and  $2 - q < l - d - 1$  to obtain

$$\int_{|z| > |x|/(2\epsilon)} |z| \Phi(z) dz \leq C \int_{|x|/(2\epsilon)}^{+\infty} r^{-l+1} r^{d-1} dr = C \left( \frac{|x|}{2\epsilon} \right)^{d-l+1} \leq C \epsilon |x|^{q-2}.$$

Therefore, (3.1) follows and we conclude that both  $K(x)$  and  $K_\epsilon(x)$  are strictly increasing in each coordinate for  $|x| > 2$ .

We now show that  $K(x)$  and  $K_\epsilon(x)$  become arbitrarily large as  $|x| \rightarrow +\infty$ . This is immediate for  $K(x)$ , since  $q > 0$ . Furthermore, since  $K(x)$  and  $\Phi_\epsilon(x)$  are nonnegative and  $\int_{\mathbb{R}^d} \Phi_\epsilon(x) dx = 1$ , we also have

$$K_\epsilon(x) \geq \int_{|y| < |x|/2} K(x-y) \Phi_\epsilon(y) dy \geq \min_{|z| \geq |x|/2} K(z) \xrightarrow{|x| \rightarrow +\infty} +\infty.$$

Thus, for both  $E$  and  $E_\epsilon$ , there exists a minimizer in  $\mathcal{P}(\mathbb{R}^d)$  [54, Theorem 3.1]. This minimizer is compactly supported when  $\epsilon > 0$  is sufficiently small, thus it belongs to  $\mathcal{P}_2(\mathbb{R}^d)$  [18, Theorem 1.4].  $\square$

*Remark 3.3.* Since  $\mathcal{P}_2(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$ , the infimum of the energy  $E$  or  $E_\epsilon$  over  $\mathcal{P}(\mathbb{R}^d)$  is less than or equal to the infimum over  $\mathcal{P}_2(\mathbb{R}^d)$ . Due to the fact that minimizers have compact support [18, Lemma 2.10], the converse holds, as well. Hence,

$$\inf_{\mu \in \mathcal{P}(\mathbb{R}^d)} E(\mu) = \inf_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} E(\mu) \quad \text{and} \quad \inf_{\mu \in \mathcal{P}(\mathbb{R}^d)} E_\epsilon(\mu) = \inf_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} E_\epsilon(\mu)$$

for all  $\epsilon > 0$  sufficiently small.

Now we prove the regularized energies  $E_\epsilon$  converge to  $E$  in the sense of  $\Gamma$ -convergence.

**Theorem 3.4** ( $\Gamma$ -convergence of regularized energies). *Suppose the energy  $E$  satisfies (E1)–(E2) and the mollifier  $\varphi$  satisfies (M1)–(M3). Then the sequence of regularized energies  $\{E_\epsilon\}_{\epsilon>0}$   $\Gamma$ -converges to the energy with respect to  $(\mathcal{P}_2(\mathbb{R}^d), d_W)$ . That is,*

- (i) (Lower semicontinuity) *For any  $\{\mu_\epsilon\}_{\epsilon>0} \subset \mathcal{P}_2(\mathbb{R}^d)$  and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  such that it holds  $\lim_{\epsilon \rightarrow 0} d_W(\mu_\epsilon, \mu) = 0$ , we have that*

$$\liminf_{\epsilon \rightarrow 0} E_\epsilon(\mu_\epsilon) \geq E(\mu).$$

- (ii) (Recovery sequence) *For any  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  there exists  $\{\nu_\epsilon\}_{\epsilon>0} \subset \mathcal{P}_2(\mathbb{R}^d)$  such that*

$$\lim_{\epsilon \rightarrow 0} d_W(\nu_\epsilon, \mu) = 0 \quad \text{and} \quad \lim_{\epsilon \rightarrow 0} E_\epsilon(\nu_\epsilon) = E(\mu).$$

*Proof.* We will prove this theorem in two steps.

Step 1 (Lower semicontinuity). By Lemma 2.5, we have  $E_\epsilon(\mu_\epsilon) = E(\mu_\epsilon * \varphi_\epsilon)$ . Furthermore,

$$(3.2) \quad d_W(\mu, \mu_\epsilon * \varphi_\epsilon) \leq d_W(\mu, \mu_\epsilon) + d_W(\mu_\epsilon, \mu_\epsilon * \varphi_\epsilon),$$

where the first term approaches zero by hypothesis and the second term approaches zero by Lemma 2.3.

As the interaction potential  $K$  is lower semicontinuous and bounded from below, the Portmanteau Theorem [60, Theorem 1.3.4] ensures the energy  $E$  is lower semicontinuous with respect to convergence in  $\mathcal{P}_2(\mathbb{R}^d)$ . Indeed, the Portmanteau Theorem states that the weak-\* convergence of  $\mu_\epsilon \times \mu_\epsilon$  to  $\mu \times \mu$  is equivalent to the fact that

$$\liminf_{\epsilon \rightarrow 0} \iint_{\mathbb{R}^n \times \mathbb{R}^n} f(x - y) d(\mu_\epsilon \times \mu_\epsilon)(x, y) \geq \iint_{\mathbb{R}^n \times \mathbb{R}^n} f(x - y) d(\mu \times \mu)(x, y)$$

when  $f$  is lower-semicontinuous and bounded from below. Hence, we obtain

$$\liminf_{\epsilon \rightarrow 0} E_\epsilon(\mu_\epsilon) = \liminf_{\epsilon \rightarrow 0} E(\mu_\epsilon * \varphi_\epsilon) \geq E(\mu).$$

Step 2 (Recovery sequence). Let  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  be arbitrary. We define a recovery sequence  $\nu_\epsilon$  for the measure  $\mu$  by using the heat kernel. Let  $\psi(x) = (4\pi)^{-d/2} e^{-|x|^2/4}$ , and define  $\psi_{\delta(\epsilon)}(x) := \delta(\epsilon)^{-d} \psi(x/\delta(\epsilon))$  where

$$(3.3) \quad \delta(\epsilon) := \epsilon^{1/2d}.$$

Clearly the function  $\psi$  satisfies the assumptions (M1)–(M3) in Subsection 2.2; hence, it is an admissible mollifier. Define the measure

$$\nu_\epsilon := \mu * \psi_{\delta(\epsilon)}.$$

Then, by Lemma 2.3,  $\nu_\epsilon \in \mathcal{P}_2(\mathbb{R}^d)$  for all  $\epsilon > 0$ , and  $d_W(\nu_\epsilon, \mu) \rightarrow 0$  as  $\epsilon \rightarrow 0$ .

We now show that  $\lim_{\epsilon \rightarrow 0} \mathbf{E}_\epsilon(\nu_\epsilon) = \mathbf{E}(\mu)$ . By Lemma 2.5,

$$(3.4) \quad \begin{aligned} |\mathbf{E}_\epsilon(\nu_\epsilon) - \mathbf{E}(\mu)| &= |\mathbf{E}(\mu * \psi_{\delta(\epsilon)} * \varphi_\epsilon) - \mathbf{E}(\mu)| = |\mathbf{E}_{\delta(\epsilon)}(\mu * \varphi_\epsilon) - \mathbf{E}(\mu)| \\ &\leq |\mathbf{E}_{\delta(\epsilon)}(\mu * \varphi_\epsilon) - \mathbf{E}_{\delta(\epsilon)}(\mu)| + |\mathbf{E}_{\delta(\epsilon)}(\mu) - \mathbf{E}(\mu)| =: I_1 + I_2. \end{aligned}$$

It suffices to show that  $I_1, I_2 \rightarrow 0$  as  $\epsilon \rightarrow 0$ . We estimate  $I_2$  first. As  $\mathbf{E}_{\delta(\epsilon)}(\mu)$  is regularized using the heat kernel  $\psi_{\delta(\epsilon)}$ , Proposition 2.6 ensures that  $\mathbf{E}_{\delta(\epsilon)}^r(\mu) \leq \mathbf{E}^r(\mu)$  for all  $\epsilon > 0$ . Hence,

$$\limsup_{\epsilon \rightarrow 0} \mathbf{E}_{\delta(\epsilon)}(\mu) \leq \mathbf{E}^r(\mu) + \limsup_{\epsilon \rightarrow 0} \mathbf{E}_{\delta(\epsilon)}^a(\mu).$$

Again, by Lemma 2.5,  $\mathbf{E}_{\delta(\epsilon)}^a(\mu) = \mathbf{E}^a(\mu * \psi_{\delta(\epsilon)})$ . Since  $d_W(\mu * \psi_{\delta(\epsilon)}, \mu) \rightarrow 0$  as  $\epsilon \rightarrow 0$  and the attractive interaction kernel  $K^a$  is a continuous function with at most quadratic growth,

$$\limsup_{\epsilon \rightarrow 0} \mathbf{E}_{\delta(\epsilon)}^a(\mu) = \limsup_{\epsilon \rightarrow 0} \mathbf{E}^a(\mu * \psi_{\delta(\epsilon)}) = \mathbf{E}^a(\mu).$$

Therefore

$$\limsup_{\epsilon \rightarrow 0} \mathbf{E}_{\delta(\epsilon)}(\mu) \leq \mathbf{E}(\mu).$$

On the other hand, by the lower semicontinuity of the energy  $\mathbf{E}$ , as proved in the first step of the proof, we get that

$$\liminf_{\epsilon \rightarrow 0} \mathbf{E}_{\delta(\epsilon)}(\mu) = \liminf_{\epsilon \rightarrow 0} \mathbf{E}(\mu * \psi_{\delta(\epsilon)}) \geq \mathbf{E}(\mu);$$

hence,  $I_2 = |\mathbf{E}_{\delta(\epsilon)}(\mu) - \mathbf{E}(\mu)| \rightarrow 0$  as  $\epsilon \rightarrow 0$ .

To estimate  $I_1$ , recall that by Remark 2.7 and 2.8 the regularized energies  $\pm \mathbf{E}_{\delta(\epsilon)}$  are  $\lambda_{\delta(\epsilon)}$ -convex with  $\lambda_{\delta(\epsilon)} = -C_\varphi \delta(\epsilon)^{-d}$  and have metric local slope  $\|\nabla K * \psi_{\delta(\epsilon)} * \mu\|_{L^2(\mu)}$ . Therefore, they satisfy the following HWI-type inequality [1, Theorem 2.4.9],

$$|\mathbf{E}_{\delta(\epsilon)}(\mu * \varphi_\epsilon) - \mathbf{E}_{\delta(\epsilon)}(\mu)| \leq 2\|\nabla K * \psi_{\delta(\epsilon)} * \mu\|_{L^2(\mu)} d_W(\mu * \varphi_\epsilon, \mu) - \frac{\lambda_{\delta(\epsilon)}}{2} d_W^2(\mu * \varphi_\epsilon, \mu).$$

Consequently, for  $\epsilon > 0$  sufficiently small, so that  $d_W(\mu * \varphi_\epsilon, \mu) \leq 1$ , we may use the bound  $\|\nabla K * \psi_{\delta(\epsilon)} * \mu\|_{L^2(\mu)} \leq C_{\psi, \mu} \delta(\epsilon)^{1-d}$  from Remark 2.8 and the definition of  $\delta(\epsilon)$  from equation (3.3) to obtain

$$I_1 = |\mathbf{E}_{\delta(\epsilon)}(\mu * \varphi_\epsilon) - \mathbf{E}_{\delta(\epsilon)}(\mu)| \leq C(\delta(\epsilon)^{1-d} + \delta(\epsilon)^{-d}) d_W(\mu * \varphi_\epsilon, \mu) \leq C\epsilon \delta(\epsilon)^{-d} = C\epsilon^{1/2}$$

Therefore  $I_1 \rightarrow 0$  as  $\epsilon \rightarrow 0$ , which completes the proof.  $\square$

*Remark 3.5.* Since the energy  $\mathbf{E}$  is lower semicontinuous with respect to weak-\* convergence of measures in  $\mathcal{P}(\mathbb{R}^d)$  and convergence in  $\mathcal{P}_2(\mathbb{R}^d)$  implies weak-\* convergence in  $\mathcal{P}(\mathbb{R}^d)$ , the conclusion of Theorem 3.4(i) is also true with respect to weak-\* convergence in  $\mathcal{P}(\mathbb{R}^d)$ .

As a result of the strong confining forces induced by the attractive part of the kernel  $K$ , we obtain the following compactness result for the sequence of energies  $\mathbf{E}_\epsilon$  in  $\mathcal{P}(\mathbb{R}^d)$ .

**Proposition 3.6** (Compactness in  $\mathcal{P}(\mathbb{R}^d)$ ). *Suppose the energy  $E$  satisfies (E1)–(E2) and the mollifier  $\varphi$  satisfies (M1)–(M3). Let  $\{\mu_\epsilon\}_{\epsilon>0} \subset \mathcal{P}(\mathbb{R}^d)$  be a sequence such that, for all  $\epsilon > 0$  sufficiently small,  $E_\epsilon(\mu_\epsilon) \leq C$  for some constant  $C > 0$ . Then  $\{\mu_\epsilon\}_{\epsilon>0}$  has a subsequence which is, up to translations, convergent with respect to the weak-\* topology in  $\mathcal{P}(\mathbb{R}^d)$ .*

*Proof.* We will use Lemma 3.1 and argue by contradiction, as in [54]. Note that as in the proof of Proposition 3.2,  $K(x)$  and  $K_\epsilon(x) \rightarrow +\infty$  as  $|x| \rightarrow +\infty$  for  $\epsilon > 0$  sufficiently small. This ensures that the energy  $E_\epsilon$  exhibits long-range confinement. We will use this fact to eliminate “vanishing” and “dichotomy” possibilities of the sequence  $\{\mu_\epsilon\}_{\epsilon>0}$ .

Suppose a subsequence of  $\{\mu_\epsilon\}_{\epsilon>0}$ , which we still index by  $\epsilon > 0$ , “vanishes” in the sense of Lemma 3.1(ii). Then for any  $\delta > 0$  and  $R > 0$ , there exists  $\epsilon_0 > 0$  such that for all  $\epsilon < \epsilon_0$  and  $x \in \mathbb{R}^d$ , we have

$$\mu_\epsilon(\mathbb{R}^d \setminus B_R(x)) \geq 1 - \delta.$$

Hence, for all  $\epsilon < \epsilon_0$ ,

$$(3.5) \quad \iint_{|x-y| \geq R} d\mu_\epsilon(x) d\mu_\epsilon(y) = \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d \setminus B_R(x)} d\mu_\epsilon(y) \right) d\mu_\epsilon(x) \geq 1 - \delta.$$

Since  $K_\epsilon(x) \rightarrow +\infty$  as  $|x| \rightarrow \infty$ , given a constant  $B > 0$ , there exists  $\tilde{R} > 0$  such that for all  $|x| \geq \tilde{R}$ , we have  $K_\epsilon(x) \geq B$ . Let  $\delta < 1/2$ , and choose  $\epsilon_0$  depending on  $\delta$  and  $\tilde{R}$  such that (3.5) holds. Then, using (3.5) and the fact that  $K_\epsilon \geq 0$  for all  $\epsilon > 0$ , we get that

$$\begin{aligned} E_\epsilon(\mu_\epsilon) &= \iint_{|x-y| < \tilde{R}} K_\epsilon(x-y) d\mu_\epsilon(x) d\mu_\epsilon(y) + \iint_{|x-y| \geq \tilde{R}} K_\epsilon(x-y) d\mu_\epsilon(x) d\mu_\epsilon(y) \\ &\geq \iint_{|x-y| \geq \tilde{R}} K_\epsilon(x-y) d\mu_\epsilon(x) d\mu_\epsilon(y) \geq (1 - \delta)B. \end{aligned}$$

This contradicts the uniform bound  $E_\epsilon(\mu_\epsilon) \leq C$  when  $B$  is sufficiently large; hence, “vanishing” does not occur.

Suppose that for a subsequence of  $\{\mu_\epsilon\}_{\epsilon>0}$ , which again we index by  $\epsilon > 0$ , “dichotomy” occurs. Then, since  $K_\epsilon \geq 0$  for all  $\epsilon > 0$ ,

$$\begin{aligned} \liminf_{\epsilon \rightarrow 0} E_\epsilon(\mu_\epsilon) &\geq \liminf_{\epsilon \rightarrow 0} \int_{\mathbb{R}^d \setminus B'_R(x_\epsilon)} \int_{B_R(x_\epsilon)} K_\epsilon(x-y) d\mu_\epsilon^1(x) d\mu_\epsilon^2(y) \\ &\geq \inf_{|x| \geq R'-R} K_\epsilon(x) \alpha(1 - \alpha), \end{aligned}$$

where  $R > 0$ , the sequence  $\{x_\epsilon\}_{\epsilon>0}$ , and the measures  $\mu_\epsilon^1$  and  $\mu_\epsilon^2$  are defined as in Lemma 3.1(iii), and  $R' > R$  is arbitrary. Thus, again using the growth of  $K_\epsilon$ , we get that

$$\liminf_{\epsilon \rightarrow 0} E_\epsilon(\mu_\epsilon) \geq +\infty,$$

contradicting the uniform bound on  $E_\epsilon(\mu_\epsilon)$ .

Therefore, Lemma 3.1 ensures the sequence  $\{\mu_\epsilon\}_{\epsilon>0}$  is tight up to a translation by a sequence  $\{y_\epsilon\}_{\epsilon>0} \subset \mathbb{R}^d$ . However, the energies  $E_\epsilon$  are translation invariant, so we can replace  $\mu_\epsilon$  by  $\mu_\epsilon(\cdot + y_\epsilon)$  which yields a tight sequence. Therefore by Prokhorov’s theorem (cf. [60, Theorem 1.3.9]), there exists a subsequence  $\{\mu_\epsilon\}_{\epsilon>0}$ , still indexed by  $\epsilon > 0$ , and a measure  $\mu \in \mathcal{P}(\mathbb{R}^d)$  such that, up to translations,  $\mu_\epsilon \rightarrow \mu$  with respect to weak-\* convergence in  $\mathcal{P}(\mathbb{R}^d)$  as  $\epsilon \rightarrow 0$ .  $\square$



Classically, if a sequence of  $\Gamma$ -convergent functionals also satisfies a compactness property, then, up to a subsequence, minimizers converge to a minimizer of the limiting functional. Though our compactness result, Proposition 3.6, is established in a weaker space  $(\mathcal{P}(\mathbb{R}^d))$  than the topology in which the energies  $\Gamma$ -converge  $(\mathcal{P}_2(\mathbb{R}^d))$ , we still obtain the following corollary on the convergence of minimizers.

**Corollary 3.7** (Convergence of minimizers). *Suppose the energy  $E$  satisfies (E1)–(E2) and the mollifier  $\varphi$  satisfies (M1)–(M3). For any  $\epsilon > 0$  sufficiently small let  $\mu_\epsilon \in \mathcal{P}_2(\mathbb{R}^d)$  be a minimizer of the energy  $E_\epsilon$ . Then there exists  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  such that, up to a subsequence and translations,  $\mu_\epsilon \rightarrow \mu$  in  $\mathcal{P}_2(\mathbb{R}^d)$  as  $\epsilon \rightarrow 0$ , and  $\mu$  minimizes the energy  $E$  over  $\mathcal{P}_2(\mathbb{R}^d)$ .*

*Proof.* Since the sequence  $\{\mu_\epsilon\}_{\epsilon>0} \subset \mathcal{P}_2(\mathbb{R}^d)$  consists of minimizers of  $E_\epsilon$ , for  $\epsilon > 0$  sufficiently small, there exists a constant  $C > 0$  so that  $E_\epsilon(\mu_\epsilon) \leq C$ . Then by Proposition 3.6, there exists a measure  $\mu \in \mathcal{P}(\mathbb{R}^d)$  such that a subsequence of  $\{\mu_\epsilon\}_{\epsilon>0}$ , which we still denote by  $\mu_\epsilon$ , weak-\* converges to  $\mu$  in  $\mathcal{P}(\mathbb{R}^d)$ .

To show that  $\mu$  minimizes  $E$  over  $\mathcal{P}_2(\mathbb{R}^d)$  we proceed in two steps. First, consider an arbitrary measure  $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ . By Theorem 3.4 (ii), there exists a sequence  $\{\nu_\epsilon\}_{\epsilon>0} \subset \mathcal{P}_2(\mathbb{R}^d)$  so that  $d_W(\nu_\epsilon, \nu) \rightarrow 0$  as  $\epsilon \rightarrow 0$  and

$$\lim_{\epsilon \rightarrow 0} E(\nu_\epsilon) = E(\nu).$$

By Theorem 3.4 (i), Remark 3.5, and the fact that the measures  $\mu_\epsilon$  are minimizers of  $E_\epsilon$  over  $\mathcal{P}_2(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$ , we obtain

$$(3.6) \quad E(\mu) \leq \liminf_{\epsilon \rightarrow 0} E_\epsilon(\mu_\epsilon) \leq \liminf_{\epsilon \rightarrow 0} E_\epsilon(\nu_\epsilon) = E(\nu).$$

Thus,  $\mu \in \mathcal{P}(\mathbb{R}^d)$  has less energy than any other measure  $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ , so by Remark 3.3,  $\mu$  minimizes  $E$  over  $\mathcal{P}(\mathbb{R}^d)$ , as well. Consequently, [18, Lemma 2.10] ensures  $\mu$  is compactly supported; hence it is in  $\mathcal{P}_2(\mathbb{R}^d)$ .

Finally, we show that in fact  $\mu_\epsilon$  converges to  $\mu$  in  $\mathcal{P}_2(\mathbb{R}^d)$ . In [18, Lemma 2.10] the authors show that

$$\text{diam}(\text{supp } \mu_\epsilon) \leq M_\epsilon$$

where  $M_\epsilon := \sqrt{d}(4r_\epsilon + (\lceil 1/m_\epsilon \rceil - 1)(4r_\epsilon + 2R_\epsilon))$ . Here

$$m_\epsilon := \frac{C - E_\epsilon(\mu_\epsilon)}{C - K_\epsilon^{\min}}$$

with  $C > 0$  so that  $E_\epsilon(\mu_\epsilon) \leq C$  for  $\epsilon > 0$  sufficiently small.  $K_\epsilon^{\min}$  denotes the absolute minimum value of the interaction kernel  $K_\epsilon$ ,  $r_\epsilon$  is chosen such that  $K_\epsilon(x) \geq C$  for all  $|x| \geq r_\epsilon$ , and  $R_\epsilon$  is the radius after which the kernel  $K_\epsilon$  is strictly increasing. We denote the corresponding quantities for the unregularized energy  $E$  by removing the subscript  $\epsilon$ .

We show that  $M_\epsilon$ , the upper bound on the diameter of the support of  $\mu_\epsilon$ , is bounded by some  $\tilde{M} > 0$  for all  $\epsilon > 0$  sufficiently small. We have  $\limsup_{\epsilon \rightarrow 0} -E_\epsilon(\mu_\epsilon) \leq -E(\mu)$  and  $K_\epsilon \rightarrow K$  uniformly on compact sets away from the origin. Thus,  $m_\epsilon \rightarrow m$  as  $\epsilon \rightarrow 0$ , so  $\lceil 1/m_\epsilon \rceil \leq \lceil 1/m \rceil + 1$  for  $\epsilon > 0$  sufficiently small.

As shown in the proof of Proposition 3.2,  $K_\epsilon$  is strictly increasing for  $|x| \geq 2$ , so we may take  $R_\epsilon = 2$ . Likewise, by estimate (3.1), there exists  $r > 2$  so that  $K_\epsilon(x) \geq C$  for  $|x| > r$ . Therefore

$$M_\epsilon \leq \tilde{M} := \sqrt{d}(4r + (\lceil 1/m \rceil)(4r + 4))$$

for all  $\epsilon > 0$  sufficiently small.

This shows that the diameter of the support of  $\mu_\epsilon$  is uniformly bounded by  $\tilde{M}$ . Consequently, the second moments of the sequence  $\mu_\epsilon$  are uniformly integrable

$$\lim_{k \rightarrow +\infty} \int_{\{|x|^2 \geq k\}} |x|^2 d\mu_\epsilon = \lim_{k \rightarrow +\infty} \int_{\{\sqrt{k} \leq |x| \leq \tilde{M}\}} |x|^2 d\mu_\epsilon = 0$$

Since  $\mu_\epsilon$  converges to  $\mu$  with respect to the weak-\* topology on  $\mathcal{P}(\mathbb{R}^d)$  and  $\mu_\epsilon$  has uniformly integrable second moments,  $\mu_\epsilon \rightarrow \mu$  in  $\mathcal{P}_2(\mathbb{R}^d)$  with respect to  $d_W$  [1, Proposition 7.1.5].  $\square$

*Remark 3.8* (Negative attraction power). The condition  $q > 0$  is a sufficient condition for Propositions 3.2 and 3.6. If  $q < 0$ , both  $K$  and  $K_\epsilon$  converge to zero as  $|x| \rightarrow \infty$ . In [54], the authors characterized the existence of minimizers for these types of potentials using the existence of a measure for which the energy is negative. This is clearly true for  $\mathbf{E}$  and  $\mathbf{E}_\epsilon$ , as one can consider the characteristic function of a sufficiently large ball  $B_R(0)$ . However, we still require  $q > 0$  for the compactness result in Proposition 3.6, and we believe that it is a necessary condition, as well.

#### 4. MORE GENERAL INTERACTION ENERGIES

Our results of sections 2 and 3 can be extended to more general nonlocal interaction energies

$$(4.1) \quad \mathbf{E}(\mu) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(x-y) d\mu(x) d\mu(y)$$

that are not necessarily repulsive-attractive and are defined via an interaction kernel  $K : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  that satisfies the following hypotheses for some fixed  $R > 1$ :

(H1)  $K$  is even, i.e.,  $K(x) = K(-x)$  for all  $x \in \mathbb{R}^d$ .

(H2)  $K \in L^1_{\text{loc}}(\mathbb{R}^d) \cap C^1(\mathbb{R}^d \setminus B_R(0))$  and  $|\nabla K(x)| \leq C(1 + |x|)$  for  $|x| > R$ .

(H3) For  $|x| > R$ ,  $K$  is strictly increasing in each coordinate and

$$\lim_{|x| \rightarrow +\infty} K(x) = +\infty.$$

(H4) There exists a function  $K^a \in C(\mathbb{R}^d)$  so that  $K^r := K - K^a$  is superharmonic.

(H5) For  $|x|$  large,  $K^r$  and  $K^a$  have at most quadratic growth, i.e.,

$$|K^a(x)| + |K^r(x)| \leq C(1 + |x|^2) \text{ for } |x| > R.$$

We briefly recall the definition of superharmonicity and some immediate properties of superharmonic functions.

**Definition 4.1.** A function  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$  is superharmonic if

$$f(x) \geq \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y) dy$$

for a.e.  $x \in \mathbb{R}^d$  and every  $r > 0$ , where  $|B_r(x)|$  denotes the volume of the ball.

*Remark 4.2* (Properties of superharmonic functions). If  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$  is superharmonic, then  $f$  is lower semicontinuous, it is bounded below on compact sets, and its distributional Laplacian satisfies  $\Delta f \leq 0$  [42, Theorem 9.3].

Interaction energies (4.1) defined via kernels satisfying assumptions (H1)–(H5) cover a wide range of applications. In particular, in any dimension  $d \geq 1$ , they include energies with  $K$  of one of the following forms:

- Attractive-repulsive power-law kernels (1.3), for  $2 - d \leq p < 0 < q \leq 2$ . (This includes Newtonian repulsion,  $p = 2 - d$ , in  $d \geq 3$ .)
- Nonsingular attractive-repulsive power-law kernels (1.3), for  $0 < p < q \leq 2$ . (This includes Newtonian repulsion,  $p = 1$ , in  $d = 1$ .)
- For  $d \geq 2$ , attractive-repulsive kernels with logarithmic repulsion,  $K(x) = |x|^q/q - \log(|x|)$  for  $0 < q \leq 2$ . (This includes Newtonian repulsion in  $d = 2$ .)
- Attractive-repulsive kernels with power-law behavior at the origin and infinity, i.e., radial kernels for which there exist  $0 < R_1 < R_2$  and  $C_1, C_2 > 0$  so that for either  $2 - d \leq p < 0 < q \leq 2$  or  $0 < p < q \leq 2$ ,

$$K(x) = C_1 |x|^p \text{ if } d \geq 1 \text{ or } K(x) = -C_1 \log(|x|) \text{ if } d \geq 2 \text{ on } \overline{B_{R_1}(0)};$$

$$K(x) \text{ is continuous on } \overline{B_{R_2}(0)} \setminus B_{R_1}(0); \text{ and,}$$

$$K(x) = C_2 |x|^q \text{ on } \mathbb{R}^d \setminus B_{R_2}(0).$$

*Remark 4.3.* Any function  $K$  satisfying the hypotheses (H1)–(H5) is locally integrable, bounded from below, and lower semicontinuous. Consequently, without loss of generality, we will assume that  $K$  is nonnegative, since the minimizers of an interaction energy  $\mathbf{E}$  with kernel  $K$  are the same as the minimizers of  $\mathbf{E}$  with kernel  $K - \inf K$ .

Our results continue to hold for  $K$  satisfying (H1)–(H5), provided that we define the regularized energies  $\mathbf{E}_\epsilon$  via compactly supported mollifiers. That is, we replace mollifier assumption (M3) by

(M3')  $\varphi$  is compactly supported in  $\mathbb{R}^d$ .

This condition eliminates possible oscillations in  $K_\epsilon$  for large values of  $|x|$ , allowing us to establish growth and monotonicity properties of the regularized kernels when  $|x|$  is sufficiently large.

We now describe how our previous results naturally extend to  $K$  satisfying (H1)–(H5). The key result in section 2 is Proposition 2.6. Assumption (H4) ensures that the potentially singular part of the kernel,  $K^r$ , is locally integrable and superharmonic. Hence, Proposition 2.6 continues to hold since, again taking  $\psi_\epsilon$  to be the heat kernel,

$$\frac{\partial}{\partial \epsilon} K_\epsilon^r = 2\psi_\epsilon * K^r * (\partial_\epsilon \psi_\epsilon) = 2\psi_\epsilon * K^r * (\Delta \psi_\epsilon) = 2\psi_\epsilon * (\Delta K^r) * \psi_\epsilon \leq 0,$$

by the superharmonicity of  $K^r$ .

The remaining results of section 2 also continue to hold for energies satisfying (H1)–(H5) and mollifiers satisfying (M1), (M2), and (M3'). In particular, (H2) ensures  $K \in L_{\text{loc}}^1(\mathbb{R}^d)$  so that Remark 2.7 holds with  $\lambda_\epsilon = C_\varphi \epsilon^{-d-2}$ , since for  $0 < \epsilon < 1$ ,

$$(4.2) \quad \|D^2 K_\epsilon\|_{L^\infty} \leq \|D^2 \varphi_\epsilon\|_{L^\infty} \int_{\text{supp } \varphi} K(x) dx \leq C_\varphi \epsilon^{-d-2} < +\infty.$$

Likewise, assumption (H2) ensures that Remark 2.8 holds. As before, we may choose a cut-off function  $\eta$  to separate  $K$  into a locally integrable component  $K^s$  near the origin and a continuously differentiable component  $K^n$  away from the origin. However, since for  $0 < \epsilon < 1$ ,

$$(4.3) \quad \|\nabla K_\epsilon^s\|_{L^\infty} \leq \|\nabla \varphi_\epsilon\|_{L^\infty} \int_{\text{supp } \varphi} K^s(x) dx \leq C_\varphi \epsilon^{-1-d},$$

we merely obtain the bound  $2\|\nabla K_\epsilon * \mu\|_{L^2(\mu)} \leq C_{\varphi, \mu} \epsilon^{-1-d}$  on the metric local slope.

Now, we describe how to extend the results from section 3. Proposition 3.2 merely requires that  $K$  and  $K_\epsilon$  are locally integrable, lower semicontinuous, strictly increasing in each coordinate outside of some ball, and become arbitrarily large as  $|x| \rightarrow +\infty$ . To see that the last two properties hold for  $K_\epsilon$ , note that for  $0 < \epsilon < 1$ ,

$$\begin{aligned} K_\epsilon(x) - K_\epsilon(y) &= \int_{\text{supp } \varphi} [K(x-z) - K(y-z)] \varphi_\epsilon(z) dz, \text{ and} \\ K_\epsilon(x) &\geq \min_{z \in \text{supp } \varphi} K(x-z). \end{aligned}$$

Consequently the properties for  $K_\epsilon$  are immediate consequences of the corresponding properties for  $K$ , which hold by (H3).

The proof of Theorem 3.4 is nearly identical for our more general energies, though we must take  $\delta(\epsilon) = \epsilon^{1/2(d+2)}$  to compensate for the inferior bounds on the convexity constant  $\lambda_\epsilon$  (4.2) and the metric slope (4.3). Finally, the proofs of Proposition 3.6 and Corollary 3.7 again use that  $K$  and  $K_\epsilon$  become arbitrarily large (uniformly in  $\epsilon$ ) as  $|x| \rightarrow +\infty$ , are strictly increasing in each coordinate outside of some ball (uniformly in  $\epsilon$ ), and are both nonnegative.

Consequently, though we choose to state our main results in sections 2 and 3 for energy functionals with repulsive-attractive power-law potentials satisfying (E1)–(E2) (see Remarks 1.1 and 1.2), these results naturally extend to a wider class of energy functionals (4.1) with  $K$  satisfying (H1)–(H5).

*Remark 4.4* (More singular kernels). The main difficulty in extending our results to interaction kernels that are more singular than the Newtonian potential at the origin (i.e.,  $-d < p < 2-d$ ) lies in the construction of a recovery sequence in the proof of Theorem 3.4. Although one could obtain a monotonicity result as in Proposition 2.6 for such kernels by regularizing them via fractional heat kernels, unfortunately, to our knowledge, in the parameter regime  $-d < p < 2-d$ , the decay estimates on fractional heat kernels do not guarantee that they have finite second moments. Therefore they are not admissible candidates in constructing our recovery sequences.

## 5. CONVERGENCE OF GRADIENT FLOWS

In this section, we apply our result on the  $\Gamma$ -convergence of regularized energies satisfying (E1)–(E2) to show that if the Wasserstein gradient flows of  $E_\epsilon$  are bounded in  $L^\infty(\mathbb{R}^d)$ , they converge to a metric space generalization of the gradient flow of  $E$ , known as a *curve of maximal slope*.

As previously mentioned, without regularization, energies  $E$  satisfying (E1)–(E2) are not convex (or  $\lambda$ -convex for  $\lambda < 0$ ). Consequently, they fall outside the scope of much of the existing theory on well-posedness of Wasserstein gradient flows [1, 22]. However, the regularized interaction energies  $E_\epsilon$  are  $\lambda_\epsilon = C_\varphi \epsilon^{-d}$  convex, so that for fixed  $\epsilon > 0$ , their Wasserstein gradient flows exist and are unique [22]. Thus, by showing these gradient flows  $\Gamma$ -converge

to the curve of maximal slope for the unregularized energy, we provide a link between the well-understood case of convex gradient flow and emerging results on the Wasserstein gradient flow of non-convex energies [2, 15, 23, 24, 26, 30, 44, 45].

We restrict our attention to the space of probability measures with bounded density

$$(5.1) \quad \mathcal{P}_{2,R}(\mathbb{R}^d) := \{\mu \in \mathcal{P}_{2,ac}(\mathbb{R}^d) : \|\mu\|_{L^\infty(\mathbb{R}^d)} \leq R\},$$

for any  $R > 0$ , due to the fact that, though  $\mathbf{E}$  is not convex (or  $\lambda$ -convex for  $\lambda < 0$ ), once it is restricted to  $\mathcal{P}_{2,R}(\mathbb{R}^d)$ , it possesses a generalized notion of convexity known as  $\omega$ -convexity, where  $\omega(x)$  is a log-Lipschitz modulus of convexity.

**Definition 5.1** (cf. [23, 30]). *Consider the modulus of convexity*

$$(5.2) \quad \omega(x) := \begin{cases} x |\log x| & \text{if } 0 \leq x \leq e^{-1-\sqrt{2}}, \\ \sqrt{x^2 + 2(1 + \sqrt{2})e^{-1-\sqrt{2}}x} & \text{if } x > e^{-1-\sqrt{2}}. \end{cases}$$

Then  $\mathbf{E} : \mathcal{P}_{2,R}(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$  is  $\omega$ -convex if, for all constant speed geodesics  $\mu_\alpha : [0, 1] \rightarrow \mathcal{P}_{2,R}(\mathbb{R}^d)$ , there exists  $C > 0$  so that

$$\mathbf{E}(\mu_\alpha) \leq (1 - \alpha)\mathbf{E}(\mu_0) + \alpha\mathbf{E}(\mu_1) + \frac{C}{2} [(1 - \alpha)\omega(\alpha^2 d_W^2(\mu_0, \mu_1)) + \alpha\omega((1 - \alpha)^2 d_W^2(\mu_0, \mu_1))].$$

This notion has been used, explicitly and implicitly, in many previous works on non-convex gradient flow [23, 24]. For our purposes,  $\omega$ -convexity allows us to define a notion of upper gradient for  $\mathbf{E}$ , which is an essential component in defining its curve of maximal slope.

Both from the perspective of energy minimization and the dynamics of the aggregation equation (1.4), the restriction to bounded densities is quite natural when  $\mathbf{E}$  is a repulsive-attractive interaction energy satisfying (E1)–(E2) with  $p = 2 - d$ . Indeed, several results in the literature support the fact that both the energy minimization and evolution take place in the space of bounded functions. For quadratic attraction (i.e., for  $q = 2$ ), the global minimizer of  $\mathbf{E}$  is the characteristic function on a ball [21, 28]. Likewise, from the dynamics point of view, if the initial data belongs to  $\mathcal{P}_2(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$  and has compact support, then it remains bounded for all time and converges to the characteristic function on the unit ball [14]. For more general attraction powers ( $0 < q \leq 2$ ), all compactly supported local minimizers are bounded [21] and smooth, compactly supported classical solutions of the aggregation equation remain bounded for all time [5, Lemma 1].

For  $p \neq 2 - d$ , the restriction to measures with bounded density is perhaps less natural, since the minimizers may concentrate mass on sets of measure zero and classical solutions to the aggregation equation can approach these steady states asymptotically [4]. However, since our proof regarding the  $\Gamma$ -convergence of gradient flows works for any choice of parameters  $2 - d \leq p < 0 < q \leq 2$ , we choose to include this regime for the sake completeness.

*Remark 5.2* (Gradient flow vs. aggregation equation). While the aggregation equation (1.4) does inform our choice of the space of measures with bounded density (5.1), the relationship between weak solutions of the aggregation equation and the Wasserstein gradient flow of the interaction energy is purely formal for the unregularized (hence nonconvex) interaction energy. There is hope this relationship can be made rigorous following the approach of Ambrosio and Serfaty [2], but we leave this for future work.

For the regularized energy, if  $\mu(t) \in \mathcal{P}_2(\mathbb{R}^d)$  is a weak solution of the aggregation equation (1.4), then it is also a Wasserstein gradient flow of  $E_\epsilon$  [1, Corollary 11.1.8]. Below, we consider the *curve of maximal slope* of  $E_\epsilon$  on the metric space of measures with bounded density  $(\mathcal{P}_{2,R}(\mathbb{R}^d), d_W)$ , and if  $\mu(t) \in \mathcal{P}_2(\mathbb{R}^d)$  is a weak solution of the aggregation equation with  $\|\mu(t)\|_\infty \leq R$ , then it is a curve of maximal slope for the weak upper gradient  $g_\epsilon(\mu) := 2\|\nabla K_\epsilon * \mu\|_{L^2(\mu)}$  on  $(\mathcal{P}_{2,R}(\mathbb{R}^d), d_W)$  [1, Theorem 11.1.3]. The reverse implication is in general false, due to the height constraint imposed by  $\mathcal{P}_{2,R}(\mathbb{R}^d)$ .

**5.1. Curves of maximal slope and  $\Gamma$ -convergence.** We now briefly recall the notion of curves of maximal slope on a complete metric space  $(\mathcal{S}, d)$ . We refer the reader to the book by Ambrosio, Gigli, and Savaré [1, Chapter 1] for further details. A curve  $u(t) : (a, b) \rightarrow \mathcal{S}$  is *2-absolutely continuous* if there exists  $m \in L^2(a, b)$  so that

$$(5.3) \quad d(u(t), u(s)) \leq \int_s^t m(r) dr \text{ for all } a < s \leq t < b.$$

For any 2-absolutely continuous curve, the limit

$$|u'(t)| = \lim_{s \rightarrow t} \frac{d(u(s), u(t))}{|s - t|}$$

exists for a.e.  $t \in (a, b)$ . Furthermore  $m(t) := |u'(t)| \in L^2(a, b)$  satisfies (5.3) and for any  $m \in L^2(a, b)$  satisfying (5.3), we have

$$|u'(t)| \leq m(t) \text{ for a.e. } t \in (a, b).$$

Given a functional  $F : \mathcal{S} \rightarrow (-\infty, +\infty]$  that is *proper*, i.e.,  $D(F) = \{u \in \mathcal{S} : F(u) < +\infty\} \neq \emptyset$ , its upper gradient is a generalization of the modulus of the gradient from Euclidean space. Specifically,  $g : \mathcal{S} \rightarrow [0, +\infty]$  is a *strong upper gradient* for  $F$  if for every 2-absolutely continuous curve  $u(t) : (a, b) \rightarrow \mathcal{S}$  the function  $g \circ u$  is measurable and

$$(5.4) \quad |F(u(t)) - F(u(s))| \leq \int_s^t g(u(r)) |u'(r)| dr \text{ for all } a < s \leq t < b.$$

One example of a strong upper gradient is given by the metric local slope

$$|\partial F|(u) := \limsup_{v \rightarrow u} \frac{(F(u) - F(v))_+}{d(u, v)},$$

when  $F$  is a  $\lambda$ -convex and lower semicontinuous functional [1, Corollary 2.4.10].

Finally, we recall the definition of a curve a maximal slope. A locally 2-absolutely continuous curve  $u : (a, b) \rightarrow \mathcal{S}$  is a *curve of maximal slope* for  $F$  with respect to the strong upper gradient  $g$  if there exists a non-increasing function  $\phi$  so that  $\phi(t) = F \circ u(t)$  for a.e.  $t \in (a, b)$  and

$$(5.5) \quad \phi'(t) \leq -\frac{1}{2}|u'(t)|^2 - \frac{1}{2}g^2(u(t)) \text{ for a.e. } t \in (a, b).$$

For all  $\epsilon > 0$ , [1, Corollary 2.4.12] ensures that if  $\mu \in D(E_\epsilon)$ , then there exists a curve of maximal slope  $\mu_\epsilon(t) : (0, +\infty) \rightarrow \mathcal{P}_{2,R}(\mathbb{R}^d)$  for  $E_\epsilon$  with respect to the strong upper gradient  $g_\epsilon(\nu) = 2\|\nabla K_\epsilon * \nu\|_{L^2(d\nu)}$  satisfying  $\mu_\epsilon(0) = \mu$ .

With these definitions in hand, we now recall a general result of Serfaty on the  $\Gamma$ -convergence of gradient flows on a metric space.



**Theorem 5.3** (cf. [53, Theorem 2]). *Let  $F_\epsilon$  and  $F$  be functionals defined on metric spaces  $(\mathcal{S}_\epsilon, d_\epsilon)$  and  $(\mathcal{S}, d)$  with strong upper gradients  $g_\epsilon$  and  $g$ , respectively. Suppose the following criteria hold:*

(i) ( $\Gamma$ -liminf convergence) *There is a notion of convergence  $S$  of  $u_\epsilon \in \mathcal{S}_\epsilon$  to  $u \in \mathcal{S}$  so that*

$$u_\epsilon \xrightarrow{S} u \quad \text{implies} \quad \liminf_{\epsilon \rightarrow 0} F_\epsilon(u_\epsilon) \geq F(u).$$

(ii) (*Lower bound on metric derivatives*) *If  $u_\epsilon(t) \xrightarrow{S} u(t)$  for  $t \in (0, T)$ , then for  $s \in [0, T)$ ,*

$$\liminf_{\epsilon \rightarrow 0} \int_0^s |u'_\epsilon|_{d_\epsilon}^2(t) dt \geq \int_0^s |u'|_d^2(t) dt.$$

(iii) (*Lower bound on slopes*) *If  $u_\epsilon \xrightarrow{S} u$ , then  $\liminf_{\epsilon \rightarrow 0} g_\epsilon(u_\epsilon) \geq g(u)$ .*

*If  $u_\epsilon(t)$  is a curve of maximal slope on  $(0, T)$  for  $F_\epsilon$  with respect to  $g_\epsilon$  satisfying*

$$u_\epsilon(t) \xrightarrow{S} u(t) \text{ for } t \in (0, T) \quad \text{and} \quad \lim_{\epsilon \rightarrow 0} F_\epsilon(u_\epsilon(0)) = F(u(0)),$$

*then  $u(t)$  is a curve of maximal slope for  $F$  with respect to  $g$  and*

$$\lim_{\epsilon \rightarrow 0} F_\epsilon(u_\epsilon(t)) = F(u(t)) \quad \text{for all } t \in [0, T),$$

$$g_\epsilon(u_\epsilon) \rightarrow g(u) \text{ and } |u'_\epsilon|_{d_\epsilon} \rightarrow |u'|_d \text{ in } L^2_{\text{loc}}(0, T).$$

*Remark 5.4.* Although the metric  $d$  induces a natural topology on  $\mathcal{S}$ , the above result admits a notion of convergence  $S$  that can be induced by a weaker topology  $\sigma$  on  $\mathcal{S}$ . Indeed,  $u_\epsilon \xrightarrow{S} u$  means that  $\pi_\epsilon(u_\epsilon) \xrightarrow{\sigma} u$  for some map  $\pi_\epsilon: \mathcal{S}_\epsilon \rightarrow \mathcal{S}$  (see [53] and [1, Remark 2.0.5] for details).

## 5.2. $\Gamma$ -convergence of the curves of maximal slope for the regularized energies.

Serfaty's result on the  $\Gamma$ -convergence of curves of maximal slope provides a powerful general framework. In practice, it can be challenging to verify conditions (i) – (iii). However, for the regularized interaction energies  $E_\epsilon$ , these conditions follow from our previous results and the following HWI-type inequality, from work by the first author, which allows us to compute the strong upper gradients.

**Lemma 5.5** (c.f. [30]). *Suppose  $\mu, \nu \in \mathcal{P}_{2,R}(\mathbb{R}^d)$  and  $E$  is an interaction energy satisfying (E1)–(E2). For  $\omega(x)$  as in equation (5.2), there exists  $C_{R,d,p,q} > 0$  so that*

$$|E(\mu) - E(\nu)| \leq 2\|\nabla K * \mu\|_{L^2(\mu)} d_W(\mu, \nu) + C_{R,d,p,q} \omega(d_W^2(\mu, \nu)).$$

With this, we show that the curves of maximal slope for the regularized energies  $\Gamma$ -converge to the curve of maximal slope of the unregularized energy.

**Theorem 5.6.** *Suppose the energy  $E$  satisfies (E1)–(E2) and the mollifier  $\varphi$  satisfies (M1)–(M3). Let  $\mu_\epsilon(t): (0, T) \rightarrow \mathcal{P}_{2,R}(\mathbb{R}^d)$  be a curve of maximal slope of  $E_\epsilon$  for the strong upper gradient  $g_\epsilon(\mu) := 2\|\nabla K_\epsilon * \mu\|_{L^2(\mu)}$  on the metric space  $(\mathcal{P}_{2,R}(\mathbb{R}^d), d_W)$ . Suppose that  $\mu_\epsilon(0)$  is well-prepared, in the sense that for some  $\mu(0) \in \mathcal{P}_{2,R}(\mathbb{R}^d)$ ,*

$$\mu_\epsilon(0) \rightarrow \mu(0) \text{ weak-}^* \text{ in } \mathcal{P}(\mathbb{R}^d) \text{ and } \lim_{\epsilon \rightarrow 0} E_\epsilon(\mu_\epsilon(0)) = E(\mu(0)).$$

Then for all  $t \in (0, T)$ ,  $\mu_\epsilon(t)$  has a weak- $*$  convergent subsequence  $\mu_\epsilon(t) \xrightarrow{\epsilon \rightarrow 0} \mu(t)$ , and  $\mu(t) : (0, T) \rightarrow \mathcal{P}_{2,R}(\mathbb{R}^d)$  is a curve of maximal slope of  $E$  for the strong upper gradient  $g(\mu) := 2\|\nabla K * \mu\|_{L^2(\mu)}$ . Furthermore, as  $\epsilon \rightarrow 0$ ,

$$\begin{aligned} E_\epsilon(\mu_\epsilon(t)) &\rightarrow E(\mu(t)) \text{ for all } t \in [0, T), \\ 2\|\nabla K_\epsilon * \mu_\epsilon\|_{L^2(\mu_\epsilon)} &\rightarrow 2\|\nabla K * \mu\|_{L^2(\mu)} \text{ in } L^2_{\text{loc}}(0, T), \\ \text{and } |\mu'_\epsilon|_{d_W} &\rightarrow |\mu'|_{d_W} \text{ in } L^2_{\text{loc}}(0, T). \end{aligned}$$

*Remark 5.7.* For any  $\mu(0) \in \mathcal{P}_{2,R}(\mathbb{R}^d)$ , there exists  $\mu_\epsilon(0) \in \mathcal{P}_{2,R}$  satisfying the conditions of the theorem: simply define  $\mu_\epsilon(0)$  by convolving  $\mu(0)$  with the heat kernel, as in Theorem 3.4(ii).

*Proof.* Let  $(\mathcal{S}_\epsilon, d_\epsilon) = (\mathcal{S}, d) = (\mathcal{P}_{2,R}(\mathbb{R}^d), d_W)$ . Note that

$$\mathcal{P}_{2,R}(\mathbb{R}^d) = \{\mu \in \mathcal{P}_2(\mathbb{R}^d) : \|\mu\|_{L^\infty(\mathbb{R}^d)} \leq R\}$$

is closed with respect to  $d_W$ , thus  $\mathcal{P}_{2,R}(\mathbb{R}^d)$  is a complete metric space. Furthermore, any  $d_W$  bounded set of  $\mathcal{P}_{2,R}(\mathbb{R}^d)$  is relatively compact with respect to weak- $*$  convergence in  $\mathcal{P}(\mathbb{R}^d)$ , and its limit points lie in  $\mathcal{P}_{2,R}(\mathbb{R}^d)$ . Given  $\mu_\epsilon \in \mathcal{S}_\epsilon$ ,  $\mu \in \mathcal{S}$  we say  $\mu_\epsilon \xrightarrow{S} \mu$  if the  $\mu_\epsilon$  converges with respect to the weak- $*$  convergence in  $\mathcal{P}(\mathbb{R}^d)$ .

We now define the strong upper gradients. For all  $\epsilon > 0$ , Remark (2.8) ensures that  $g_\epsilon(\nu) = 2\|\nabla K_\epsilon * \nu\|_{L^2(\nu)}$  is a strong upper gradient of  $E_\epsilon$  on  $\mathcal{P}_2(\mathbb{R}^d)$ , hence on  $\mathcal{P}_{2,R}(\mathbb{R}^d)$ . Next, we show that  $g(\nu) = 2\|\nabla K * \nu\|_{L^2(\nu)}$  is a strong upper gradient of  $E$ . Throughout, we use the fact that if  $\mu \in L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ , then  $K * \mu(x)$  is continuously differentiable and  $\nabla(K * \mu) = (\nabla K) * \mu$ .

Suppose  $\nu(t) : (a, b) \rightarrow \mathcal{P}_{2,R}(\mathbb{R}^d)$  is an 2-absolutely continuous curve. The function  $t \mapsto g(\nu(t))$  is measurable, since it is given by the composition of measurable functions. By Lemma 5.5,

$$(5.6) \quad |E(\nu(t)) - E(\nu(s))| \leq 2\|\nabla K * \nu(t)\|_{L^2(\nu(t))} d_W(\nu(t), \nu(s)) + C_{R,d,p,q} \omega(d_W^2(\nu(t), \nu(s))).$$

As in Remark 2.8, we estimate  $\|\nabla K * \nu(t)\|_{L^2(\nu(t))}$  by breaking  $\nabla K$  into its singular and nonsingular parts,

$$\begin{aligned} \|\nabla K * \nu(t)\|_{L^2(\nu(t))} &\leq \|\nabla K^n * \nu(t)\|_{L^2(\nu(t))} + \|\nabla K^s * \nu(t)\|_{L^2(\nu(t))} \\ &\leq \left( \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \nabla K^n(x-y) d\nu(y, t) \right|^2 d\nu(x, t) \right)^{1/2} + \sqrt{R} \|\nabla K^s * \nu(t)\|_{L^2(\mathbb{R}^d)} \\ &\leq \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} |\nabla K^n(x-y)|^2 d\nu(x, t) \right)^{1/2} d\nu(y, t) + \sqrt{R} \|\nabla K^s\|_{L^1(\mathbb{R}^d)} \|\nu(t)\|_{L^2(\mathbb{R}^d)}. \end{aligned}$$

Since  $\nabla K^n$  has at most linear growth, the first term is bounded by the second moments of  $\nu(t)$ , which are uniformly bounded for  $t \in (a, b)$  since  $\{\nu(t)\}_{t \in (a,b)}$  is bounded with respect to  $d_W$ . The second term is uniformly bounded since  $\|\nu(t)\|_{L^1} = 1$  and  $\|\nu(t)\|_{L^\infty} \leq R$ .

By the absolute continuity of  $\nu(t)$ ,  $d_W(\nu(t), \nu(s)) \leq \int_s^t |\nu'(r)| dr$  for  $|\nu'| \in L^2(a, b)$ . Combining this with (5.6) ensures that for almost every  $t$

$$\begin{aligned} \left| \frac{d}{dt} E(\nu(t)) \right| &\leq \lim_{s \rightarrow t} \frac{|E(\nu(t)) - E(\nu(s))|}{|t - s|} \\ &\leq \lim_{s \rightarrow t} 2 \|\nabla K * \nu(t)\|_{L^2(\nu(t))} \frac{d_W(\nu(t), \nu(s))}{|t - s|} + C_{R,d,p,q} \frac{\omega(d_W^2(\nu(t), \nu(s)))}{|t - s|} \\ &\leq 2 \|\nabla K * \nu(t)\|_{L^2(\nu(t))} |\nu'(t)|, \end{aligned}$$

where in the last line we use that, for  $\omega(x)$  defined by (5.2),  $\omega(x^2) \leq 4x\omega(x)$  and  $\lim_{x \rightarrow 0} \omega(x) = 0$ . Integrating the above inequality, we conclude,

$$|E(\nu(t)) - E(\nu(s))| = \left| \int_s^t \frac{d}{dr} E(\nu(r)) dr \right| \leq \int_s^t 2 \|\nabla K * \nu(r)\|_{L^2(\nu(r))} |\nu'(r)| dr.$$

Thus  $g(\nu) = 2 \|\nabla K * \nu\|_{L^2(\nu)}$  is a strong upper gradient of  $E$ .

Now we show that for all  $t \in (0, T)$ ,  $\mu_\epsilon(t) \xrightarrow{\epsilon \rightarrow 0} \mu(t) \in \mathcal{P}_{2,R}(\mathbb{R}^d)$ , with respect to weak-\* convergence of probability measures. Since  $\mu_\epsilon(t) : [0, \infty) \rightarrow \mathcal{P}_2(\mathbb{R}^d)$  is a curve of maximal slope for the strong upper gradient  $g_\epsilon(\mu) = 2 \|\nabla K_\epsilon * \mu\|_{L^2(\mu)}$ , applying Cauchy's inequality to (5.4) and comparing with (5.5), we obtain (see [1, Remark 1.3.3])

$$(5.7) \quad E_\epsilon(\mu_\epsilon(s)) - E_\epsilon(\mu_\epsilon(t)) = \int_s^t |\mu'_\epsilon|^2(r) dr$$

for all  $0 \leq s \leq t < \infty$ . By definition of the metric local slope and the fact that  $E_\epsilon(\mu_\epsilon(t)) \geq 0$ ,

$$d_W^2(\mu_\epsilon(0), \mu_\epsilon(t)) \leq \left( \int_0^t |\mu'_\epsilon|(r) dr \right)^2 \leq t \int_0^t |\mu'_\epsilon|^2(t) dt \leq T E_\epsilon(\mu_\epsilon(0))$$

for all  $t \in (0, T)$ . Since  $\lim_{\epsilon \rightarrow 0} E_\epsilon(\mu_\epsilon(0)) = E(\mu(0))$ , the right hand side is uniformly bounded for  $\epsilon$  sufficiently small. Therefore,  $\{\mu_\epsilon(t)\}_{\epsilon > 0}$  is uniformly bounded in  $\mathcal{P}_{2,R}(\mathbb{R}^d)$ , and, up to a subsequence,  $\mu_\epsilon(t) \xrightarrow{\epsilon \rightarrow 0} \mu(t) \in \mathcal{P}_{2,R}(\mathbb{R}^d)$  with respect to weak-\* convergence in  $\mathcal{P}(\mathbb{R}^d)$ .

It remains to verify criteria (i) – (iii) of Theorem (5.3) to conclude that  $\mu(t)$  is a curve of maximal slope of  $E$  for  $g(\mu)$ , and the corresponding energies, strong upper gradients and metric local slopes converge as  $\epsilon \rightarrow 0$ .

Criterion (i) follows immediately from Theorem 3.4, part (i) and Remark 3.5. To prove (ii), we assume without loss of generality that there exists  $0 \leq C < +\infty$  so that

$$C = \liminf_{\epsilon \rightarrow 0} \int_0^s |\mu'_\epsilon|^2(t) dt.$$

Choose a subsequence  $|\tilde{\mu}'_\epsilon|(t)$  so that  $\lim_{\epsilon \rightarrow 0} \int_0^s |\tilde{\mu}'_\epsilon|^2(t) dt = C$ . Then  $|\tilde{\mu}'_\epsilon|(t)$  is bounded in  $L^2(0, s)$  so, up to a further subsequence, it is weakly convergent to some  $v(t) \in L^2(0, s)$ . Consequently, for any  $0 \leq s_0 \leq s_1 \leq s$ ,

$$\lim_{\epsilon \rightarrow 0} \int_{s_0}^{s_1} |\tilde{\mu}'_\epsilon|(t) dt = \int_{s_0}^{s_1} v(t) dt.$$

By taking limits in the definition of the metric derivative and using the lower semicontinuity of  $d_W$  with respect to weak-\* convergence,

$$d_W(\mu_\epsilon(s_0), \mu_\epsilon(s_1)) \leq \int_{s_0}^{s_1} |\mu'_\epsilon|(t) dt \quad \text{yields} \quad d_W(\mu(s_0), \mu(s_1)) \leq \int_{s_0}^{s_1} v(t) dt.$$

By [1, Theorem 1.1.2], this implies that  $|\mu'| (t) \leq v(t)$  for a.e.  $t \in (0, s)$ . Thus, by the lower semicontinuity of the  $L^2(0, s)$  norm with respect to weak convergence,

$$\liminf_{\epsilon \rightarrow 0} \int_0^s |\mu'_\epsilon|^2(t) dt = \lim_{\epsilon \rightarrow 0} \int_0^s |\tilde{\mu}'_\epsilon|^2(t) dt \geq \int_0^s v(t)^2 dt \geq \int_0^s |\mu'|^2(t) dt.$$

Finally, we turn to (iii). We assume without loss of generality that

$$C = \liminf_{\epsilon \rightarrow 0} g_\epsilon(\mu_\epsilon),$$

for some  $0 \leq C < +\infty$ . Choose a subsequence  $g_\epsilon(\mu_\epsilon)$  so that  $\lim_{\epsilon \rightarrow 0} g_\epsilon(\mu_\epsilon) = C$ . Since  $\|\mu_\epsilon(t)\|_{L^\infty(\mathbb{R}^d)} \leq R$  for all  $\epsilon > 0$  and all  $t \in (0, T)$ , there exists a further subsequence so that  $\mu_\epsilon$  converges to some limit  $\nu$  in the weak-\* topology of  $L^\infty(\mathbb{R}^d)$ . Since  $\mu_\epsilon$  also converges to  $\mu$  with respect to the weak-\* topology of  $\mathcal{P}(\mathbb{R}^d)$ , for all  $f \in C_c^\infty(\mathbb{R}^d)$ ,

$$\int_{\mathbb{R}^d} f(x) \nu(x) dx = \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^d} f(x) \mu_\epsilon(x) dx = \int_{\mathbb{R}^d} f(x) \mu(x) dx;$$

hence,  $\nu = \mu$ .

By [1, Theorem 5.4.4] it suffices to show

$$\lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^d} f(x) (\nabla K_\epsilon * \mu_\epsilon)(x) d\mu_\epsilon(x) = \int_{\mathbb{R}^d} f(x) (\nabla K * \mu)(x) d\mu(x) \quad \text{for all } f \in C_c^\infty(\mathbb{R}^d).$$

Then using the convexity and the lower semicontinuity of the function  $|\cdot|^2$  along with the weak-\* convergence of  $\mu_\epsilon$  yields the result.

Following a similar argument as in Lemma 2.5 and defining  $\tilde{\mu}_\epsilon = \varphi_\epsilon * \mu_\epsilon * \varphi_\epsilon$ , we have

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} f(x) (\nabla K_\epsilon * \mu_\epsilon)(x) d\mu_\epsilon(x) - \int_{\mathbb{R}^d} f(x) (\nabla K * \mu)(x) d\mu(x) \right| \\ &= \left| \int_{\mathbb{R}^d} f(x) (\nabla K * \tilde{\mu}_\epsilon)(x) d\mu_\epsilon(x) - \int_{\mathbb{R}^d} f(x) (\nabla K * \mu)(x) d\mu(x) \right| \\ &\leq \left| \int_{\mathbb{R}^d} f(x) [(\nabla K * \tilde{\mu}_\epsilon)(x) - (\nabla K * \mu)(x)] d\mu_\epsilon(x) \right| \\ &\quad + \left| \int_{\mathbb{R}^d} f(x) (\nabla K * \mu)(x) d\mu_\epsilon(x) - \int_{\mathbb{R}^d} f(x) (\nabla K * \mu)(x) d\mu(x) \right| \\ &=: A_\epsilon + B_\epsilon. \end{aligned}$$

Since  $\mu_\epsilon \rightarrow \mu$  weak-\* in  $\mathcal{P}(\mathbb{R}^d)$  and  $f(x) (\nabla K * \mu)(x)$  is bounded and continuous,  $\lim_{\epsilon \rightarrow 0} B_\epsilon = 0$ .

It remains to show  $\lim_{\epsilon \rightarrow 0} A_\epsilon = 0$ . First, note that  $\tilde{\mu}_\epsilon \rightarrow \mu$  in the weak-\* topology of  $L^\infty(\mathbb{R}^d)$  as  $\epsilon \rightarrow 0$ . Indeed, for any  $f \in L^1(\mathbb{R}^d)$ , we have that

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} f(x) d\tilde{\mu}_\epsilon(x) - \int_{\mathbb{R}^d} f(x) d\mu(x) \right| = \left| \int_{\mathbb{R}^d} \varphi_\epsilon * f * \varphi_\epsilon(x) d\mu_\epsilon(x) - \int_{\mathbb{R}^d} f(x) d\mu(x) \right| \\ &\leq R \|\varphi_\epsilon * f * \varphi_\epsilon - f\|_{L^1} + \left| \int_{\mathbb{R}^d} f(x) d\mu_\epsilon(x) - \int_{\mathbb{R}^d} f(x) d\mu(x) \right|, \end{aligned}$$

where both terms approach zero as  $\epsilon \rightarrow 0$ . Returning to  $A_\epsilon$ ,

$$\begin{aligned} A_\epsilon &\leq \int_{\mathbb{R}^d} |f(x)(\nabla K * (\tilde{\mu}_\epsilon - \mu))(x)| |\mu_\epsilon(x)| dx \\ &\leq R \|f\|_{L^\infty(\mathbb{R}^d)} \int_{\text{supp } f} |\nabla K * (\tilde{\mu}_\epsilon - \mu)(x)| dx. \end{aligned}$$

Since the integrand has at most linear growth, it is bounded on the compact set  $\text{supp } f$ , and we may apply the dominated convergence theorem, provided the integrand converges pointwise.

When  $0 < q \leq 1$ ,  $\nabla K$  is the sum of a bounded continuous function and an integrable function, and since  $\tilde{\mu}_\epsilon \xrightarrow{\epsilon \rightarrow 0} \mu$  in both weak-\* probability and weak-\*  $L^\infty(\mathbb{R}^d)$ , the integrand converges for each  $x$ . On the other hand, when  $1 < q \leq 2$ , it suffices to show that  $\nabla K$  is the sum of a continuous function, which is uniformly integrable with respect to  $\tilde{\mu}_\epsilon$ , and an integrable function [1, Lemma 5.1.7]. In particular, is enough to show that  $|x|^{q-1}$  is uniformly integrable with respect to  $\tilde{\mu}_\epsilon$ . Since  $|x| \geq k$  implies that  $|x|/k \geq 1$  we have that

$$\lim_{k \rightarrow \infty} \int_{|x| \geq k} |x|^{q-1} d\tilde{\mu}_\epsilon \leq \lim_{k \rightarrow \infty} \frac{1}{k} \int_{|x| \geq k} |x|^q d\tilde{\mu}_\epsilon \leq \lim_{k \rightarrow \infty} \frac{1}{k} E(\tilde{\mu}_\epsilon) = \lim_{k \rightarrow \infty} \frac{1}{k} E_\epsilon(\mu_\epsilon).$$

Since  $\mu_\epsilon(t)$  is a curve of maximal slope for  $E_\epsilon$ , (5.7) gives  $E_\epsilon(\mu_\epsilon(t)) \leq E_\epsilon(\mu_\epsilon(0))$  for all  $t \in (0, T)$ . The well-preparedness of the initial data gives  $E_\epsilon(\mu_\epsilon(0)) < C$ , hence the result follows.  $\square$

*Remark 5.8.* The extension of Theorem 5.6 to energies defined via kernels satisfying the more general conditions (H1)–(H5) from section 4 requires also a generalization of the technical result of Lemma 5.5, which plays a crucial role in the proof of the convergence theorem. As such a generalization is beyond the scope of our current study we leave this to future studies.

**Acknowledgments.** This project was initiated during the Thematic Program on Variational Problems in Physics, Economics and Geometry at the Fields Institute for Research in Mathematical Sciences. The authors would like to thank the organizers and the Institute for their hospitality and support during their visits. The authors would also like to thank Andrea Bertozzi, Eric Carlen, Andres Contreras, Inwon Kim, and Xavier Lamy for useful discussions, along with the anonymous reviewers for their detailed reading and valuable comments.

## REFERENCES

- [1] L. AMBROSIO, N. GIGLI, AND G. SAVARÉ, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, second ed., 2008.
- [2] L. AMBROSIO AND S. SERFATY, *A gradient flow approach to an evolution problem arising in superconductivity*, Comm. Pure Appl. Math., 61 (2008), pp. 1495–1539.
- [3] C. ANDERSON AND C. GREENGARD, *On vortex methods*, SIAM J. Numer. Anal., 22 (1985), pp. 413–440.
- [4] D. BALAGUÉ, J. CARRILLO, T. LAURENT, AND G. RAOUL, *Nonlocal interactions by repulsive-attractive potentials: Radial ins/stability*, Phys. D., 260 (2013), pp. 5–25.
- [5] D. BALAGUÉ, J.A. CARRILLO, AND Y. YAO, *Confinement for repulsive-attractive kernels*, Discrete Contin. Dyn. Syst. Ser. B, 19 (2014), pp. 1227–1248.
- [6] D. BALAGUÉ, J. A. CARRILLO, T. LAURENT, AND G. RAOUL, *Dimensionality of local minimizers of the interaction energy*, Arch. Ration. Mech. Anal., 209 (2013), pp. 1055–1088.
- [7] J.T. BEALE AND A. MAJDA, *Vortex methods. I. Convergence in three dimensions*, Math. Comp., 39 (1982), pp. 1–27.
- [8] ———, *Vortex methods. II. Higher order accuracy in two and three dimensions*, Math. Comp., 39 (1982), pp. 29–52.

- [9] D. BENEDETTO, E. CAGLIOTI, J. A. CARRILLO, AND M. PULVIRENTI, *A non-Maxwellian steady distribution for one-dimensional granular media*, J. Statist. Phys., 91 (1998), pp. 979–990.
- [10] A.J. BERNOFF AND C.M. TOPAZ, *A primer of swarm equilibria*, SIAM J. Appl. Dyn. Syst., 10 (2011), pp. 212–250.
- [11] A.L. BERTOZZI AND J. BRANDMAN, *Finite-time blow-up of  $L^\infty$ -weak solutions of an aggregation equation*, Commun. Math. Sci., 8 (2010), pp. 45–65.
- [12] A.L. BERTOZZI, J.A. CARRILLO, AND T. LAURENT, *Blow-up in multidimensional aggregation equations with mildly singular interaction kernels*, Nonlinearity, 22 (2009), pp. 683–710.
- [13] A.L. BERTOZZI, J.B. GARNETT, AND T. LAURENT, *Characterization of radially symmetric finite time blowup in multidimensional aggregation equations*, SIAM J. Math. Anal., 44 (2012), pp. 651–681.
- [14] A.L. BERTOZZI, T. LAURENT, AND F. LÉGER, *Aggregation and spreading via the Newtonian potential: the dynamics of patch solutions*, Math. Models Methods Appl. Sci., 22 (2012), pp. 1140005, 39.
- [15] A.L. BERTOZZI, T. LAURENT, AND J. ROSADO,  *$L^p$  theory for the multidimensional aggregation equation*, Comm. Pure Appl. Math., 64 (2011), pp. 45–83.
- [16] A.L. BERTOZZI, H. SUN, T. KOLOKOLNIKOV, D. UMINSKY, AND J. VON BRECHT, *Ring patterns and their bifurcations in a nonlocal model of biological swarms*, Comm. Math. Sci., 13 (2015), pp. 955–985. Special issue in honor of George Papanicolaou’s birthday.
- [17] A. BLANCHET, E.A. CARLEN, AND J.A. CARRILLO, *Functional inequalities, thick tails and asymptotics for the critical mass Patlak-Keller-Segel model*, J. Funct. Anal., 262 (2012), pp. 2142–2230.
- [18] J.A. CAÑIZO, J.A. CARRILLO, AND F.S. PATACCHINI, *Existence of compactly supported global minimisers for the interaction energy*, Arch. Ration. Mech. Anal., 217 (2015), pp. 1197–1217.
- [19] J.A. CARRILLO, M. CHIPOT, AND Y. HUANG, *On global minimizers of repulsive-attractive power-law interaction energies*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 372 (2014), pp. 20130399, 13.
- [20] J.A. CARRILLO, Y.-P. CHOI, AND M. HAURAY, *The derivation of swarming models: mean-field limit and Wasserstein distances*, Collective Dynamics from Bacteria to Crowds: An Excursion Through Modeling, Analysis and Simulation Series, CISM International Centre for Mechanical Sciences, 553 (2014), pp. 1–46.
- [21] J.A. CARRILLO, M.G. DELGADINO, AND A. MELLET, *Regularity of local minimizers of the interaction energy via obstacle problems*, Preprint, 2014.
- [22] J.A. CARRILLO, M. DI FRANCESCO, A. FIGALLI, T. LAURENT, AND D. SLEPČEV, *Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations*, Duke Math. J., 156 (2011), pp. 229–271.
- [23] J.A. CARRILLO, S. LISINI, AND E. MAININI, *Uniqueness for Keller-Segel-type chemotaxis models*, Discrete Contin. Dyn. Syst., 34 (2014), pp. 1319–1338.
- [24] J.A. CARRILLO, R.J. MCCANN, AND C. VILLANI, *Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates*, Rev. Mat. Iberoamericana, 19 (2003), pp. 971–1018.
- [25] ———, *Contractions in the 2-Wasserstein length space and thermalization of granular media*, Arch. Ration. Mech. Anal., 179 (2006), pp. 217–263.
- [26] J.A. CARRILLO AND J. ROSADO, *Uniqueness of bounded solutions to aggregation equations by optimal transport methods*, in European Congress of Mathematics, Eur. Math. Soc., Zürich, 2010, pp. 3–16.
- [27] D. CHAFAÏ, N. GOZLAN, AND P.-A. ZITT, *First-order global asymptotics for confined particles with singular pair repulsion*, Ann. Appl. Probab., 24 (2014), pp. 2371–2413.
- [28] R. CHOKSI, R.C. FETECAU, AND I. TOPALOGLU, *On minimizers of interaction functionals with competing attractive and repulsive potentials*, Ann. Inst. H. Poincaré Anal. Non Linéaire, (2015). to appear.
- [29] Y.-L. CHUANG, Y.R. HUANG, M.R. D’ORSOGNA, AND A.L. BERTOZZI, *Multi-vehicle flocking: scalability of cooperative control algorithms using pairwise potentials*, IEEE International Conference on Robotics and Automation, (2007), pp. 2292–2299.
- [30] K. CRAIG, *Conditions for well-posedness of non-convex gradient flow in the Wasserstien metric*. in progress, 2015.
- [31] K. CRAIG AND A. BERTOZZI, *A blob method for the aggregation equation*, Math. Comp., (2015). to appear.
- [32] H. DONG, *The aggregation equation with power-law kernels: ill-posedness, mass concentration and similarity solutions*, Comm. Math. Phys., 304 (2011), pp. 649–664.



- [33] J.P.K. DOYE, D.J. WALES, AND R.S. BERRY, *The effect of the range of the potential on the structures of clusters*, J. Chem. Phys., 103 (1995), pp. 4234–4249.
- [34] K. FELLNER AND G. RAOUL, *Stable stationary states of non-local interaction equations*, Math. Models Methods Appl. Sci., 20 (2010), pp. 2267–2291.
- [35] R.C. FETECAU AND Y. HUANG, *Equilibria of biological aggregations with nonlocal repulsive-attractive interactions*, Phys. D, 260 (2013), pp. 49–64.
- [36] R.C. FETECAU, Y. HUANG, AND T. KOLOKOLNIKOV, *Swarm dynamics and equilibria for a nonlocal aggregation model*, Nonlinearity, 24 (2011), pp. 2681–2716.
- [37] C.R. GIVENS AND R.M. SHORTT, *A class of Wasserstein metrics for probability distributions*, Michigan Math. J., 31 (1984), pp. 231–240.
- [38] M.F. HAGAN AND D. CHANDLER, *Dynamic pathways for viral capsid assembly*, Biophysical Journal, 91 (2006), pp. 42–54.
- [39] Y. HUANG AND A.L. BERTOZZI, *Self-similar blowup solutions to an aggregation equation in  $\mathbf{R}^n$* , SIAM J. Appl. Math., 70 (2010), pp. 2582–2603.
- [40] T. KOLOKOLNIKOV, H. SUN, D. UMINSKY, AND A.L. BERTOZZI, *Stability of ring patterns arising from two-dimensional particle interactions*, Phys. Rev. E, 84 (2011).
- [41] H. LI AND G. TOSCANI, *Long-time asymptotics of kinetic models of granular flows*, Arch. Ration. Mech. Anal., 172 (2004), pp. 407–428.
- [42] E.H. LIEB AND M. LOSS, *Analysis*, vol. 14 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2001.
- [43] P.-L. LIONS, *The concentration-compactness principle in the calculus of variations. The locally compact case. I*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 1 (1984), pp. 109–145.
- [44] E. MAININI, *A global uniqueness result for an evolution problem arising in superconductivity*, Boll. Unione Mat. Ital. (9), 2 (2009), pp. 509–528.
- [45] ———, *Well-posedness for a mean field model of Ginzburg-Landau vortices with opposite degrees*, NoDEA Nonlinear Differential Equations Appl., 19 (2012), pp. 133–158.
- [46] R.J. MCCANN, *Stable rotating binary stars and fluid in a tube*, Houston J. Math., 32 (2006), pp. 603–631.
- [47] A. MOGILNER AND L. EDELSTEIN-KESHET, *A non-local model for a swarm*, J. Math. Biol., 38 (1999), pp. 534–570.
- [48] A. MOGILNER, L. EDELSTEIN-KESHET, L. BENT, AND A. SPIROS, *Mutual interactions, potentials, and individual distance in a social aggregation*, J. Math. Biol., 47 (2003), pp. 353–389.
- [49] L. PEREA, G. GÓMEZ, AND P. ELOSEGUI, *Extension of the Cucker–Smale control law to space flight formations*, AIAA J. of Guidance, Control, and Dynamics, 32 (2009), pp. 527–537.
- [50] D. PETZ AND F. HIAI, *Logarithmic energy as an entropy functional*, in Advances in differential equations and mathematical physics (Atlanta, GA, 1997), vol. 217 of Contemp. Math., Amer. Math. Soc., Providence, RI, 1998, pp. 205–221.
- [51] F. POUPAUD, *Diagonal defect measures, adhesion dynamics and Euler equation*, Methods Appl. Anal., 9 (2002), pp. 533–561.
- [52] M.C. RECHTSMAN, F.H. STILLINGER, AND S. TORQUATO, *Optimized interactions for targeted self-assembly: application to a honeycomb lattice*, Phys. Rev. Lett., 95 (2005).
- [53] S. SERFATY, *Gamma-convergence of gradient flows on Hilbert and metric spaces and applications*, Discrete Contin. Dyn. Syst., 31 (2011), pp. 1427–1451.
- [54] R. SIMIONE, D. SLEPČEV, AND I. TOPALOGLU, *Existence of ground states of nonlocal-interaction energies*, J. Stat. Phys., 159 (2015), pp. 972–986.
- [55] E.M. STEIN, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970.
- [56] M. STRUWE, *Variational Methods: Applications to Nonlinear Partial Differential Equations and Hamiltonian Systems*, Springer-Verlag, Berlin, third ed., 2000.
- [57] H. SUN, D. UMINSKY, AND A.L. BERTOZZI, *Stability and clustering of self-similar solutions of aggregation equations*, J. Math. Phys., 53 (2012), pp. 115610, 18.
- [58] C.M. TOPAZ AND A.L. BERTOZZI, *Swarming patterns in a two-dimensional kinematic model for biological groups*, SIAM J. Appl. Math., 65 (2004), pp. 152–174.
- [59] C.M. TOPAZ, A.L. BERTOZZI, AND M.A. LEWIS, *A nonlocal continuum model for biological aggregation*, Bull. Math. Biol., 68 (2006), pp. 1601–1623.

- [60] A.W. VAN DER VAART AND J. WELLNER, *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer Series in Statistics, Springer, 1996.
- [61] C. VILLANI, *Topics in optimal transportation*, vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003.
- [62] D.J. WALES, *Energy landscapes of clusters bound by short-ranged potentials*, Chem. Eur. J. Chem. Phys., 11 (2010), pp. 2491–2494.
- [63] Y. YAO AND A.L. BERTOZZI, *Blow-up dynamics for the aggregation equation with degenerate diffusion*, Phys. D, 260 (2013), pp. 77–89.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SANTA BARBARA, CA  
*E-mail address:* `kcraig@math.ucsb.edu`

DEPARTMENT OF MATHEMATICS AND STATISTICS, MCMASTER UNIVERSITY, HAMILTON, ON  
*E-mail address:* `itopalog@math.mcmaster.ca`